

# **Advancing the Protein-Catalyzed Capture Agent Technology to New Frontiers**

Thesis by

Amy Michelle McCarthy

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

The Caltech logo is rendered in a large, bold, orange sans-serif font. The letters are thick and closely spaced, with a modern, clean design.

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, CA

2018

(Defended May 21, 2018)

© 2018

Amy Michelle McCarthy

ORCID: 0000-0003-3456-0383

All rights reserved except where otherwise noted

*To my family and my fiancé Michael*

## Acknowledgements

In my six years at Caltech there has been a village of people supporting me as I complete my doctorate. First of all I want to thank my advisor Professor James Heath for being such a great advisor for the last four years. You have been very supportive both in terms of research suggestions and in securing research funding. You also allowed me to explore many different veins of research so that I could hone my research skills in a variety of areas, and I got experience with filing patents as a result. I intend to channel your penchant for picking up new skills/knowledge by finding good mentors as I start my postdoc. I would also like to thank my thesis committee members Dennis Dougherty, Jonas Peters, and Mark Davis for all of their advice and encouragement during our yearly meetings. Thirdly, my thanks also go out to my very first research advisor, my undergraduate research advisor Professor Richmond Sarpong. Thank you for letting a sophomore who was still taking organic chemistry conduct research in your lab. I really enjoyed my 2.5 years in your group, and I am very grateful for all the advice and encouragement that you have given me during my time at Caltech.

My time in the Heath group would not be the same without the fabulous members of the Heath lab. I feel great gratitude towards all the members of “team malaria” Dr. Arundhati Nag, Dr. Samir Das, Dr. Aiko Umeda, Dr. David Bunck, and (now a Dr.!) JingXin Liang. Thank you so much for taking me under your collective wings when I switched into the Heath group and teaching me how to prepare and test protein-catalyzed capture agents. I particularly miss being able to walk down the hall to talk phys org/organic synthesis with the Drs. Arundhati and Samir, but I’m so excited that the two of you have embarked on your independent careers. Dr. Jungwoo Kim was another



awesome mentor as we worked together on the B-RAP technology. Thank you for spending all those hours teaching me how to fabricate DNA barcodes. I wish you the best as you continue in your postdoc. Thank you to Alice Hsu, Yapeng Su, and Fan Liu for the many conversations that we had in our shared office. And thank you, Fan for feeding me random snacks while I was working on my paper manuscript/exit props and getting annoyed at Microsoft Word.

Even though we all spread out in different directions after our time together at UC Berkeley, several of you managed to keep in touch and support me all along in graduate school. Reyu Sakakibara, thanks so much for continuing to keep in touch while we worked on our Ph.D.s on opposite coasts. All of the texts, letters, phone calls, and packages from my favorite lab partner made a huge difference in my time at Caltech. Jeffrey Zhang thank you for continuing to check in and challenging me to make sure that I remember to take breaks from lab on occasion. And for introducing me to my fiancé; I look forward to seeing you at my wedding this June! Weruche Okaru, thank you for making time to text/call me and to hang out with me so much when you moved down to Southern California. I will miss our regular meet ups when I leave. I'd also like to thank some awesome former grad students/postdocs from Cal who kept tabs on me and gave me advice/encouragement when I needed it: Dr. Jessica Herrick, Professor Dr. Jessica Kisunzu, Dr. Zach Hallberg, Professor Dr. Sidney Wilkerson-Hill, Dr. Daniel Fischer, and Professor Dr. André Isaacs.

There have also been a lot of people at Caltech who have been a great support network for me over the last six years. Agnes Tong and Allison Ross thank you for keeping us grad students organized and for always having your office doors open for grad students.

Katrine Museth and Elyse Garlock thank you for being such an awesome lab manager and lab administrator, respectively, in the Heath lab and making sure that everything ran smoothly. We could not have conducted research without both of you! I am also grateful to the members of the Caltech Knit Happens group for many evenings of stitching and dishing as well as a variety of fun group outings over the years. You guys encouraged me to keep up with my crafting and decompress even when things were really busy in lab. Moreover, a big thank you to Dr. Alysia Mayes, Dr. Nadia Herrera, and Dr. Chinenye (Chinny) Idigo for mentoring me during my early years at Caltech and for being such great friends. I miss our pilgrimages to Mother Moos Creamery and random meet ups at Chinny's. Watching the three of you complete your Ph.D.s and go on to jobs/postdocs definitely encouraged me to keep soldiering on towards the light at the end of the tunnel! I would also like to express my profound gratitude to my awesome allergist Dr. Reyneiro Castro for doing such a great job treating my allergies and asthma when everything flared up my third year. Without your excellent care this thesis would have never been written. Thank you also to my awesome roommates over the years Rachel Ford, Jamie Rankin, and Elizabeth (Liz) Holman. We were random matches that ended up being a great support system for each other as we all worked towards our dissertations. Jamie and Liz I'll especially miss our culinary experiments over the years. I wish you all the best as you wrap up your degrees. And thank you as well to Bart Plovie, Ben Walsh, Eric Woelki, Strawberry Shortcake, and Jacob Vanderburgh for playing Age of Empires II HD with me and listening to me talk about grad school for the last couple of years. Our irregular gaming sessions the last year helped me unwind after lab.

Another aspect of my support network was my awesome family, and I'd especially like to thank my parents who agreed to put my two littlest siblings, then 11 and 14, on a plane twice a year so that I could have them hang out with me for a few days. That made the transition from living with seven other people to one other person much easier especially in the first year. A big and heartfelt thank you to my siblings Frank, Elizabeth, Dannielle, Esther, and Sarah, for their support and for regularly visiting me at Caltech. I'd also like to thank my church family at Crosslife Community Church for looking after me for the last six years, especially Laurie Preddy who prayed for me daily and the Mark family who made sure that I always had a ride to church.

Finally I would like to thank my other half, my fiancé Michael, for being a rock that I could lean on for the last 4.5 years. Thank you for cheering with me during the high points and lifting me up during the low points of grad school. Thank you for cleaning my apartment and making me delicious meals when you visit. Thank you for listening to me prattle on about my research all the time and for making me so very loved and lucky. I eagerly await our wedding in a few weeks and look forward to taking on life together with you.

## Abstract

Protein-catalyzed capture (PCC) agents are a nascent synthetic aptamer technology that was first disclosed in 2009. In addition to reviewing the different classes of peptide-based aptamers in chapter 1, this thesis records efforts to advance the PCC technology in two ways. First, in chapter 2 the development of a barcoded-rapid assay platform (B-RAP) technology enables the parallel analysis of up to fifteen PCC agents at once as well as dramatically shortening the time required to characterize the binding affinity for a pool of ligands from weeks to a couple of days. Secondly, the capture agent technology was utilized to target difficult proteins. Kirsten rat sarcoma (KRas) protein is a GTPase that acts as a light switch for several important cellular signaling pathways. Oncogenic variants of KRas are responsible for driving roughly 20-25% of all cancers, but KRas is considered “undruggable” from a small molecule targeting point of view. We report the identification of PCC ligands that bind to conserved allosteric switches on KRas and inhibit the protein’s GTPase enzymatic activity. The biomarker *Plasmodium falciparum* Histidine Rich Protein II (HRP2) presents an unusual challenge as it is a highly variable, unstructured and sticky protein. In chapter 3 we report on efforts to develop low nM binding capture agents against highly prevalent epitopes of HRP2, and the use of medicinal chemistry optimization to prepare structurally related variants of the lead capture agent for probing the structure-activity relationship and how it affects binding to HRP2.

## Published Content and Contributions

\* = equal contribution

Chapter 2 is comprises a manuscript currently under review at Analytical Chemistry.

A full patent has also been filed on the B-RAP technology.

“An Allosteric Inhibitor of KRas Identified Using a Barcoded Rapid Assay Platform”

**McCarthy, A. M.**<sup>\*</sup>; Kim, J.<sup>\*</sup>; Museth, K. A.; Henning, R. K.; Heath, J. E.; Winson, E.; Oh, J. J.; Heath, J. R. *Anal. Chem.* **2018**, *In Review*.

A.M.M. conceived research ideas, assisted in the preparation of microfluidic technologies, and performed measurements and the subsequent data analysis.

Heath, J. R.; **McCarthy, A. M.**; Kim, J. “Barcoded Rapid Assay Platform for Efficient Analysis of Candidate Molecules and Methods of Making and Using the Platform”, U.S. Patent Appl., 15942351, **2018**.

A.M.M. conceived research ideas, assisted in the preparation of microfluidic technologies, and performed measurements and the subsequent data analysis.

## Table of Contents

<b>Acknowledgements</b>	iii
<b>Abstract</b>	viii
<b>Published Content and Contribution</b>	ix
<b>Table of Contents</b>	x
<b>List of Figures and Tables</b>	xiv
<b>Abbreviations and Acronyms</b>	xxi
<b>Chapter 1: Synthetic Antibodies and Their Applications</b>	1
Abstract	1
Section 1.1-Introduction	2
Section 1.2-Affimers	3
Section 1.3-Peptides	6
Section 1.4-Protein Catalyzed Capture Agents	8
Section 1.5-Peptoids	10
Section 1.6-Analysis of the Synthetic Aptamer Classes	13
Section 1.7-Theme of the Thesis	17
Section 1.8-References	17
<b>Chapter 2: Identification of an Allosteric Inhibitor of KRas Protein Using a Barcoded Rapid Assay Platform</b>	27
Abstract	27
Section 2.1-Introduction	28
Section 2.2-Methods	31
2.2.1-Preparation of the Barcode Rapid Assay Platform	31

2.2.2-KRas Protein Expression and Purification	33
2.2.3-Preparation of Switch I and Switch II SynEps and Scrambled SynEps	33
2.2.4-Library Preparation and <i>In-situ</i> Library Click Screen	33
2.2.5-Surface Immunofluorescent Assays on the Barcoded Rapid Assay Platform	34
2.2.6-Measuring the Effect of the Allosteric Ligands on the Intrinsic KRas GTPase Activity.	35
Section 2.3-Results and Discussion	36
2.3.1-Optimizing B-RAP Technology Assay Conditions	36
2.3.2-Validation of the B-RAP Technology	37
2.3.3-Measuring the EC <sub>50</sub> of the Allosteric Binding PCC Ligands	38
2.3.4-Testing the Allosteric Ligands as Inhibitors of KRas GTPase Activity	40
Section 2.4-Conclusion	41
Section 2.5-References	42
Appendix 2A-Materials and Methods	46
List of Chemicals and Reagents	47
DNA Barcode Chip Patterning and Validation	51
Barcode Validation	52
Expression of Cysteine-Modified Streptavidin (SAC) Protein	53
Isolation of SAC Inclusion Bodies from <i>E. coli</i> Cells	53
Refolding and Purification of SAC Protein	54

Preparation of SAC-DNA Conjugates	54
Biotin Binding Test of DESL Set	56
WT KRas Expression and Purification	57
<i>In-Situ</i> Library Screen Preclear/Anti-Screen	58
<i>In-Situ</i> Library Screen Product/Target Screens	59
Peptide Synthesis Protocols	60
WT KRas Binding Curves Using the B-RAP Technology	60
WT KRas Binding Curves Using the Multi-Well ELISA Technology	62
Testing Allosteric Ligands for Inhibition of Intrinsic KRas Protein	
GTPase Activity	63
References	64
 Appendix 2B-Supplemental Figures	 65
Appendix 2C-Supplemental Tables	89
 <b>Chapter 3: Development of Epitope-Targeting Protein Catalyzed Capture</b>	
 <b>Agents Against Histidine Rich Protein II</b>	 98
Abstract	98
Section 3.1-Introduction	99
Section 3.2-Methods	101
3.2.1-Preparation of Peptides	101
3.2.2-Enzyme-Linked Immunosorbent Assay Analysis of Peptides	102
3.2.3-Linker Library Synthesis	102
3.2.4-Linker Library Screen Pre-Clear	103



3.2.5-Linker Library Product Screens	103
Section 3.3-Results and Discussion	103
3.3.1-Alanine Scan of Sequence Hits	103
3.3.2-Second-Generation Ligand Synthesis and Measurement of HRP2 Binding Affinity	105
3.3.3-Further Optimization of Ligand Hit <b>hL3V2</b>	107
3.3.4-Developing PCC Agent Biligands	110
Section 3.4-Conclusion	112
Section 3.5-References	112
Appendix 3A-Materials and Methods	114
Materials	115
General Procedures for the Preparation of Ligands	118
Overnight ELISA Procedure	121
Room Temperature ELISA Procedure	122
Linker Library Pre-Clear	123
Linker Library Product Screen	125
References	125
Appendix 3B-Supplemental Figures	126
Appendix 3C-Supplemental Tables	165
References	174

## List of Figures, Schemes, and Tables

### Chapter 1

#### Figures

Figure 1.1- Representations of each affimer class.	3
Figure 1.2- A representative schematic of the outcome of a yeast two-hybrid assay.	5
Figure 1.3- A demonstration of a combinatorial split-and-mix protocol.	7
Figure 1.4- An example of a cyclic peptide closed by a (1,4)-triazole linkage.	8
Figure 1.5- Ranking the different synthetic aptamer classes in terms of their similarities to naturally occurring biological motifs.	13

#### Schemes

Scheme 1.1- A graphical flow-chart of the <i>in-situ</i> library click screen to identify PCC agents against a SynEp.	9
Scheme 1.2-The submonomer synthesis method for the preparation of peptoids.	10
Scheme 1.3- A protocol to directly analyze the binding affinity of library screen hits directly.	12

#### Tables

Table 1.1- Scaffold used to develop affimer reagents.	4
Table 1.2- A comparison of structure and preparation of antibodies and synthetic peptide aptamers.	14
Table 1.3- Comparison of physical properties of antibodies	

and synthetic aptamers. 15

## Chapter 2

### **Figures**

Figure 2.1- Overview of the B-RAP Technology.	29
Figure 2.2- Identification of SynEps for the dual epitope in-situ click screen, and the resulting PCC candidate hit sequences.	36
Figure 2.3- Full binding curves for L1-L9 and the corresponding EC <sub>50</sub> values.	38
Figure 2.4- Measuring the inhibition of the GTPase activity of KRas by Ligand <b>L2</b> .	40
Figure A2B.1-High throughput patterning of DNA barcode chips.	67
Figure A2B.2-DNA barcode chip layout and validation.	68
Figure A2B.3-Comparison of barcode quality across an entire microchip.	69
Figure A2B.4-Probing barcode stability.	69
Figure A2B.5-Representative SAC-DNA FPLC trace.	70
Figure A2B.6-FPLC trace of wild-type KRas protein.	70
Figure A2B.7-Biotin binding evaluation of the full SAC-DNA set with a Biotin* probe on the barcoded rapid assay platform.	70
Figure A2B.8-The characterization data for SynEp1.	71
Figure A2B.9-The characterization data for scrambled SynEp1.	72
Figure A2B.10-The characterization data for SynEp2.	73
Figure A2B.11-The characterization data for scrambled SynEp2.	74
Figure A2B.12-The characterization data for Ligand L1.	75

Figure A2B.13-The characterization data for Ligand L2.	76
Figure A2B.14-The characterization data for Ligand L3.	77
Figure A2B.15-The characterization data for Ligand L4.	78
Figure A2B.16-The characterization data for Ligand L5.	79
Figure A2B.17-The characterization data for Ligand L6.	80
Figure A2B.18-The characterization data for Ligand L7.	81
Figure A2B.19-The characterization data for Ligand L8.	82
Figure A2B.20-The characterization data for Ligand L9.	83
Figure A2B.21-Single point Immunofluorescent Assay using the B-RAP technology.	84
Figure A2B.22-Statistical validation of the B-RAP Technology.	84
Figure A2B.23-A comparison of the average line-scan with an averaged data-block extraction.	85
Figure A2B.24-Pixel extraction of full barcode lanes in a given barcode.	86
Figure A2B.25-KRas binding curves without dummy ligand extraction.	87
Figure A2B.26-KRas protein binding curves using 96-well plate ELISAs.	87
Figure A2B.27-The wild-type KRas protein GTPase activity curves.	88
Figure A2B.28-Controls for the L2 GTPase inhibition curve.	88

## Tables

Table 2.1- Comparison of the capacity, reagent quantities used, and assay times for multi-well ELISA plates relative to the barcoded rapid assay platform.	41
Table A2C.1-A comparison of the various steps associated with the	

Solution Loading Method (reported here) and the previously reported Tubing Method for preparing microfluidic flow-patterned ssDNA barcodes.	90
Table A2C.2-Table of ssDNA sequences used for SAC-DNA conjugation and DNA barcode patterning.	91
Table A2C.3-Analysis of the degradation in barcode quality during extended storage.	92
Table A2C.4-Synthetic epitopes and PCC ligands characterization table.	93
Table A2C.5-A pixel-by-pixel analysis of variance along a barcode lane.	95
Table A2C.6-The %CV data for the ELISAs.	95
Table A2C.7-Calculated p-values for the pixel-by-pixel full-lane and centroid scatterplots.	96
Table A2C.8-The goodness of fit measurements for the allosteric KRas binding curves.	97

### Chapter 3

#### **Figures**

Figure 3.1-The single point ELISA results for each alanine scan set.	104
Figure 3.2-Testing the second-generation ligands.	106
Figure 3.3-Evaluating the stapled and ring expanded triazole ligands.	109
Figure 3.4-Structure of the PCC anchors used in the linker study.	111
Figure A3B.1-Characterization spectra for ligand hL1A1.	127
Figure A3B.2-Representative characterization spectra for ligand hL1A2.	128
Figure A3B.3-Representative characterization spectra for ligand hL1A3.	129

Figure A3B.4-Characterization spectra for ligand hL1A4.	130
Figure A3B.5-Representative characterization spectra for ligand hL1A5.	131
Figure A3B.6-Characterization spectra for ligand hL2.	132
Figure A3B.7-Representative characterization spectra for ligand hL2A1.	133
Figure A3B.8-MALDI spectra for ligand hL2A2.	134
Figure A3B.9-MALDI spectra for ligand hL2A3.	135
Figure A3B.10-Characterization spectra for ligand hL2A4.	136
Figure A3B.11-Characterization spectra for ligand hL2A5.	137
Figure A3B.12-Representative characterization spectra for ligand hL3A1.	138
Figure A3B.13-Representative characterization spectra for ligand hL3A2.	139
Figure A3B.14-Characterization spectra for ligand hL3A3.	140
Figure A3B.15-Characterization spectra for ligand hL3A4.	141
Figure A3B.16-Characterization spectra for ligand hL3A5.	142
Figure A3B.17-Characterization spectra for ligand hL1D3.	143
Figure A3B.18-Characterization spectra for ligand hL1E3.	144
Figure A3B.19-Characterization spectra for ligand hL2fF1.	145
Figure A3B.20-Representative characterization spectra for ligand hL2f*F1.	146
Figure A3B.21-Characterization spectra for ligand hL2fF2.	147
Figure A3B.22-Representative characterization spectra for ligand hL2f*F2.	148

Figure A3B.23-Characterization spectra for ligand hL3V2.	149
Figure A3B.24-Characterization spectra for ligand hL3.	150
Figure A3B.25-Characterization spectra for ligand hL3f*F3.	151
Figure A3B.26-Characterization spectra for ligand hL3fF4.	152
Figure A3B.27-Characterization spectra for hL3f*F4.	153
Figure A3B.28-Characterization spectra for ligand hL3fF5.	154
Figure A3B.29-Characterization spectra for ligand hL3f*F5.	155
Figure A3B.30-Representative characterization spectra for ligand Ak-hL3G1V3.	156
Figure A3B.31-Characterization spectra for ligand Ak-hL3V2G6.	157
Figure A3B.32-Representative characterization spectra for ligand Ak'-hL3G1V3.	158
Figure A3B.33-Characterization spectra for ligand Ak'-hL3V2G6.	159
Figure A3B.34-MALDI spectra for ligand Tz-hL3G1V3.	160
Figure A3B.35-Characterization spectra for ligand Tz-hL3V2G6.	161
Figure A3B.36-Representative characterization spectra for C-term ligand.	162
Figure A3B.37-Characterization spectra for ligand DNBCtermH8.	163
Figure A3B.38-Structure of the spontaneously formed heterodimer in the click reaction.	164

## Tables

Table 3.1-The fourteen unique epitopes found in HRP2.	100
Table 3.2-The prevalence of the different HRP2 epitopes in patient	

samples from endemic areas.	101
Table A3C.1-Alanine scan sets.	166
Table A3C.2-Generation two ligands.	166
Table A3C.3-Alternative variants of <b>hL3V3</b> prepared.	166
Table A3C.4-Linker hits from JXL1 vs. N-term linker screen.	166
Table A3C.5-Linker hits from DNB vs. C-term linker screen.	167
Table A3C.6-Linker hits from the DNB vs. N-term linker screen.	167
Table A3C.7-Linker hits from JXL1 vs. DNB linker screen.	167
Table A3C.8-Linker hits from JXL1 vs. C-term linker screen.	167
Table A3C.9-Ligand characterization data.	168



## List of Abbreviations and Acronyms

A, Ala (L-Alanine)

ab (antibody)

AD (Activating domain)

ADC (Antibody-drug conjugate)

Aib (aminoisobutyric acid)

Akt1

p-Akt2

E17K Akt1

ATP (adenosine triphosphate)

Az4, <sup>4-N3</sup>K ((4-azido)-L-lysine)

BCIP (5-bromo-4-chloro-3-indolylphosphate toluidine salt)

BD (Binding domain)

Biotin\* (Biotin-polyadenylate-Cy3)

βME (beta-mercaptoethanol)

Boc (tert-butyloxy carbonyl)

BoNT (Botulinium neurotoxin)

B-RAP (barcoded rapid assay platform)

BSA (bovine serum albumin)

BS3 (bis(sulfosuccinimidyl)suberate)

<sup>t</sup>Bu (tert-butyl group)

%CV (percent coefficient of variation)

CAPPIA (Cell Array Protein-Protein Interaction Assay)

CDK2 (Cyclin-dependent kinase 2)

CDK4 (Cyclin-dependent kinase 4)

CRP (C-Reactive Protein)

CuAAC (Copper (I)-catalyzed azide-alkyne cycloaddition)

D, Asp (L-Aspartic acid/L-Aspartate)

Da (Dalton)

DCM (dichloromethane)

DEAL (DNA-encoded antibody library)

DIC (diisopropyl carbodiimide)

DMSO (dimethylsulfoxide)

DNA (deoxyribonucleic acid)

E, Glu (L-Glutamic acid/L-Glutamate)

*E. coli* (*Escherichia coli*)

EC<sub>50</sub> (half maximal effective concentration)

ELISA (enzyme-linked immunosorbent assay)

Et<sub>2</sub>O (diethyl ether)

F<sub>532</sub> (relative fluorescence units at 532 nm excitation wavelength)

F<sub>635</sub> (relative fluorescence units at 635 nm excitation wavelength)

<sup>3,4</sup>-F<sup>F</sup> (L-(3, 4-difluoro)-Phenylalanine)

<sup>4</sup>F<sup>F</sup> (L-(4-fluoro)-Phenylalanine)

Fmoc (fluorenyl methyloxycarbonyl)

Fmoc-Oct-OH (Fmoc-(S)- $\alpha$ -Me-2'-(7-octenyl)-alanine)

Fmoc-Pen-OH (Fmoc-(R)- $\alpha$ -Me-2'-(4-pentenyl)-alanine)

FPLC (fast protein liquid chromatography)

f.u. (fluorescence units)

G, Gly (Glycine)

GDP (guanosine diphosphate)

GFP (Green Fluorescent Protein)

Guan•HCl (guanidinium chloride)

GST (Glutathione S-transferase)

GTP (guanosine triphosphate)

H, His (L-Histidine)

HATU (1-[Bis (dimethylamino) methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid  
hexafluorophosphate)

H<sub>2</sub>O (water)

H<sub>2</sub>SO<sub>4(aq)</sub> (sulfuric acid)

HER4 (human epidermal growth factor receptor 4)

HPLC (high performance liquid chromatography)

HRP (Horseradish Peroxidase)

HRP2 (Histidine Rich Protein II)

I, Ile (L-Isoleucine)

IB (inclusion body)

IC<sub>50</sub> (half-maximum inhibitory concentration)

IFA (immunofluorescent assay)

IgG (Immunoglobulin G)

IL-17F (Interleukin 17-F)

IL8 (Interleukin 8)

1,6-IPTG (isopropyl  $\beta$ -D-1-thiogalactopyranoside)

$K_D$  (dissociation constant)

KRas (Kristen Rat Sarcoma)

L, Leu (L-leucine)

l (D-Leucine)

Lys, K (L-lysine)

MALDI-TOF (matrix assisted laser desorption ionization time-of-flight)

MAPPIT (Mammalian Protein-Protein Interaction)

mIgG (Mouse Immunoglobulin G)

MeCN (acetonitrile)

Met, M (L-methionine)

$MgCl_2 \cdot 5H_2O$  (magnesium chloride pentahydrate)

MHPH (maleimide 6-hydrazino-nicotinamide)

$MnCl_2$  (manganese (II) chloride)

mRNA (messenger RNA)

MS (mass spectrometry)

MWCO (molecular weight cut-off)

NACS (nucleic acid cell sorting)

NaCl (sodium chloride)

$NaHCO_3$  (sodium bicarbonate)

NaOAc (sodium acetate)

$NH_4OAc$  (ammonium acetate)

NBT (nitro-blue tetrazolium chloride)

NLS (Nuclear localizing sequence)

NMP (N-methyl-2-pyrrolidone)

OBOC (one bead one compound)

P, Pro (L-proline)

p, (D-proline)

PBS (phosphate buffered saline)

PCC (protein catalyzed capture)

PD1 (Programmed cell death protein 1)

PDMS (polydimethylsulfoxide)

PEG<sub>n</sub> (polyethylene glycol)

pf (*Plasmodium falciparum*)

pfLDH (*Plasmodium falciparum* lactate dehydrogenase)

pxLDH (pan-*Plasmodium* lactate dehydrogenase)

PLL (poly-L-lysine)

Pra,  $\alpha$ -propargylGly (( $\alpha$ -propargyl)-glycine)

$4\text{-N}^3$ Pro ((4-azido)-L-proline)

Q, Gln (L-Glutamine)

Ras (rat sarcoma)

RCM (ring closing metathesis)

RNA (ribonucleic acid)

RT (room temperature)

S-4FB (N-succinimidyl-4-formylbenzaldehyde)

SAC (cysteine-modified streptavidin)

SDS (sodium dodecyl sulfate)

Semi-prep (semi-preparative)

SOD1 (Superoxide dismutase 1)

SPPS (solid phase peptide synthesis)

SPS (solid-phase synthesis)

SQM (Stefin A Quadruple Mutant Protein)

ssDNA (single-stranded DNA)

STM (Stefin A Triple Mutant Protein)

SynEp (synthetic epitope)

T, Thr (threonine)

TBS (tris-buffered saline)

TCEP (tris-(2-carboxyethyl)phosphine hydrochloride)

TESH (triethylsilane)

TF (Transcription Factor)

TFA (trifluoroacetic acid)

TNF- $\alpha$  (Tumor Necrosis Factor alpha)

Tris (tris(hydroxymethyl)aminomethane)

Tris•HCl (tris(hydroxymethyl)aminomethane)

TrxA (Thioredoxin A protein)

Tween20 (polysorbate 20)

V, Val (L-Valine)

Y, Tyr (L-Tyrosine)

## **Chapter 1**

### **Synthetic Antibody Analogues and their Applications**

Antibodies are a form of biological aptamer that have been utilized in the development of therapeutics and diagnostics. However, antibodies are a large biomolecule that is not perfectly optimized for these applications. This fueled the rise of synthetic aptamer classes to replace antibodies including peptide-based aptamer classes. In this review we look at the development and application of the peptide-based synthetic aptamer classes affimers, free-peptides, protein-catalyzed capture agents, and peptoids.

# Synthetic Antibody Analogues and their Applications

## Section 1.1-Introduction

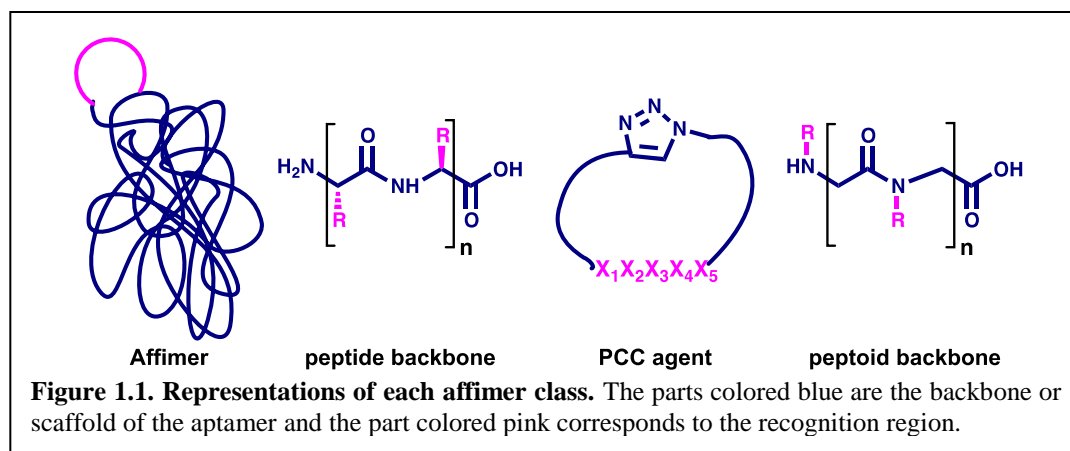
Biological aptamers such as antibodies form an integral component of immune systems and have a variety of jobs *in vivo* including distinguishing cell types, detecting foreign invaders, and triggering allergic reactions.<sup>1</sup> Antibodies have also been utilized in the development of cancer treatments and microarray diagnostic kits. For example, antibody drug conjugates (ADCs) link a cancer drug to an antibody raised against specific cell type in order to deliver the drug to cancer cells bound by the antibody.<sup>2</sup> Secondly, immunotherapy treatments employ antibodies such as anti-PD1 to block the interactions between a T-cell and a tumor cell that prevent the T-cell from killing the tumor cell.<sup>3,4</sup> Additionally, the evolution of a patient's proteomics over time can be tracked via pull-down of various proteins using antibody microarrays such as the DNA Encoded Antibody Library (DEAL) technology.<sup>5-7</sup>

Antibodies are valuable biological tools, but they possess characteristics that present a challenge for their utilization. For example, interspecies antibodies are immunogenic and cannot be used in humans. Secondly, antibodies are the result of injecting an animal, usually a mammal, and extracting the resulting antibodies from their bloodstream. The use of animals to generate antibodies makes the antibody quality dependent on the animal's health and which species of animals is being used. Antibodies are also large proteins with the average IgG antibody weighing ~150 kDa.<sup>8</sup> As a result, antibodies have poor cellular penetration and are susceptible to thermal and proteasome degradation. Furthermore, antibodies are poorly optimized for direct surface immobilization, as this can affect their structure, which limits their usefulness in directly-patterned microarrays.



Some disclosed workarounds to the above issues include humanizing antibodies,<sup>9</sup> producing known antibodies in transformed cells,<sup>10</sup> and substituting single-domain antibody fragments for full-antibodies.<sup>11,12</sup> However, an alternative approach would be to develop a synthetic replacement, which mimics the selectivity and binding affinity of an antibody, but is better suited for biomedical applications. Synthetic aptamers<sup>13</sup> designed as antibody surrogates can be prepared from various biomolecules including nucleic acids<sup>14</sup> and peptides.<sup>15</sup> The peptide-based aptamer class includes the subclasses of affimers or peptamers,<sup>16</sup> free peptides, protein-catalyzed capture (PCC) agents,<sup>17</sup> and peptoids<sup>18</sup> (for representative structures of each see Figure 1.1). The remainder of this review will focus on these classes of peptide-based aptamers.

## Section 1.2-Affimers



Peptamers consist of a modified small protein (~15 kDa) that serves as a scaffold to one or more peptide loops introduced into its structure by the use of restriction enzymes. The first peptamers were generated on the modified Thioredoxin A (TrxA) protein scaffold,<sup>19</sup> with multiple scaffolds reported since (Table 1.1).<sup>19–25</sup> The inserted peptides range in length from roughly 8-20 residues, which mimics the size of the antibody recognition

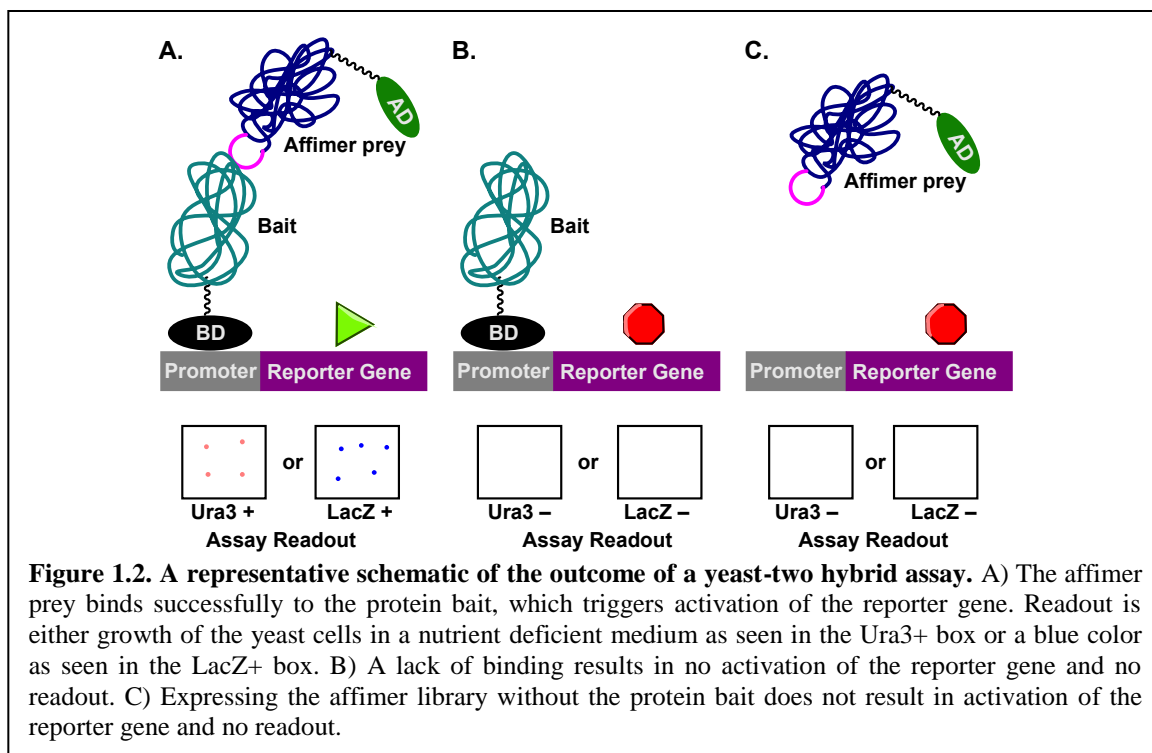
region while minimizing the number of defective peptamers in a library as a result of a random stop codon present in the peptide sequence.<sup>26</sup>

**Table 1.1. Scaffolds used to develop affimer reagents.**

Scaffold	Ref.
TrxA	19
STM	20
GFP	21
Affibodies	22
Anticalins	23
Adhiron	24
SQM	25

Peptamers libraries are generated by the preparation of plasmids containing the protein scaffold with the inserted random peptide sequences. The use of two orthogonal restriction enzyme sites to add each peptide sequence into the scaffold yields libraries with the greatest number of functional affimers, as the peptide sequence can only be inserted in a single direction.<sup>26</sup> A typical peptamer library ranges in size from  $10^6$  to  $10^{12}$ , with library sizes  $>10^9$  favored. The resulting plasmids are then transformed into eukaryotic yeast or mammalian cells in preparation for the library screen. Screening in eukaryotic cells ensures proper folding of the target proteins and tests the peptamers in conditions that mimic conditions found in the human body. Scale up production of affimer hits can occur in either *E. coli* or eukaryotic cell-based expression systems.

There are several different methods to screen peptamer libraries. One of the most common screening methods is the yeast two-hybrid system (Figure 1.2).<sup>26,27</sup> This screen exploits the discovery that the discrete DNA-binding domain (BD) and activating domain (AD) of certain transcription factors can still induce expression if they are fused to different molecules, but still in close albeit with reduced expression efficiency.<sup>27</sup> In the yeast two-hybrid system the TF's BD is fused to the target protein to produce the "bait", and the "prey" consists of the TF's AD-domain fused to both a nuclear localizing sequence (NLS) and unique members of the peptamer library. The gene under the TF's control is chosen carefully such that the screen's readout is either the growth of cells in



media lacking essential nutrients (Ura3, Leu2, Ade2, His3, etc.) or the development of color due to enzymatic activity (LacZ).<sup>28</sup> Screens conducted in mammalian cells utilize a similar “bait and prey” interaction approach with different readouts: mammalian protein-protein interaction trap (MAPPIT, cytokine activation readout)<sup>29,30</sup> and cell array protein-protein interaction array (CAPPIA, fluorescent readout).<sup>31</sup> Mammalian-cell based screens are even better mimics of *in vivo* conditions than yeast cells, which can be important when developing affimers for *in vivo* applications. For libraries with the Adhiron scaffold, screening involves iterative bio-panning of phage-displayed peptamers.<sup>24</sup> After the library screen cells containing a “hit” are lysed, and the collected plasmids are sequenced before scale-up expression of the peptamers via standard protocols.

The last twenty-two years has seen affimers utilized as functional effectors on proteins<sup>26,32</sup> with the biotech company Avacta working to commercialize the affimer technology. Affimers, especially affimers based on the STM or Adhiron scaffolds, have

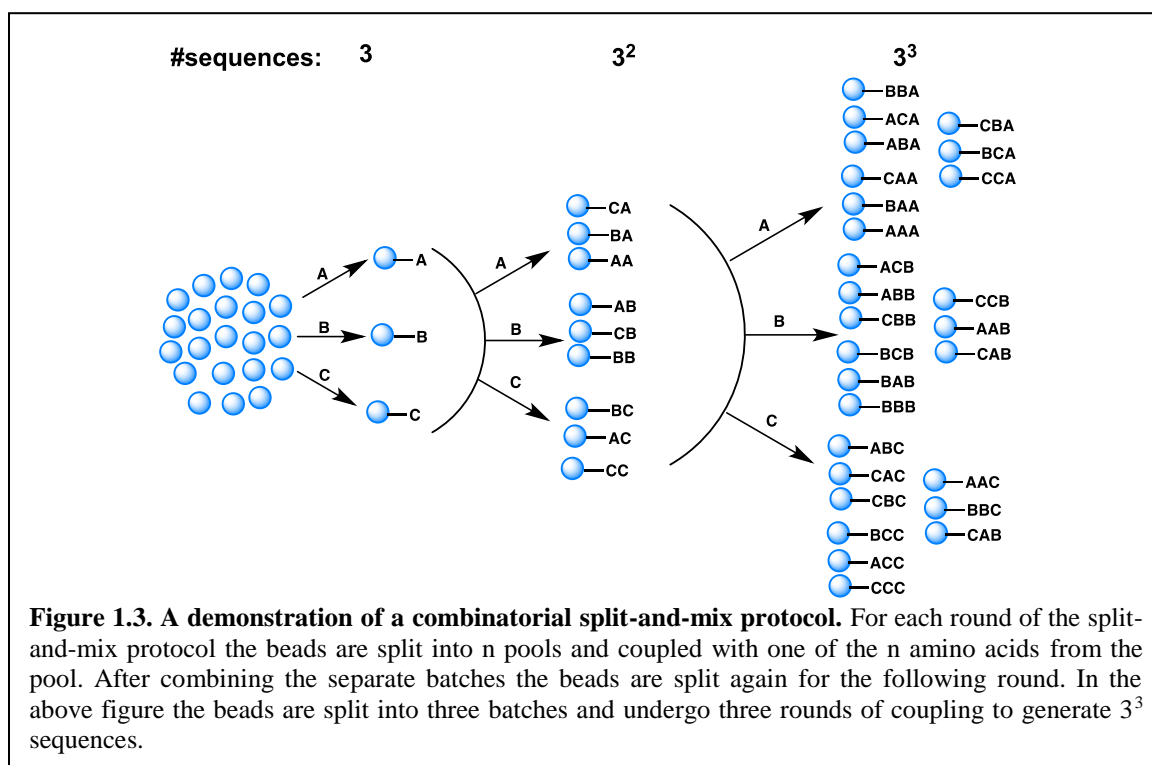
also been engineered such that they can be attached to microelectrodes to form aptasensors.<sup>33–39</sup> The aptasensors can form a microarray for the detection of cancer biomarkers or for proteomic analysis of a serum sample. Peptasensor microarrays can be readily produced using dot-printing production and other microfabrication protocols.

### Section 1.3-Peptides

Peptides consist of a short polypeptide chain (<30 residues) and can be prepared as either a linear polypeptide or as a stapled (cyclic) macrocyclic peptide, which will be focused on in this review. Stapled peptides have increased affinity and proteolytic stability relative to linear peptides likely as a result of their constrained structure. Cyclization reactions include head-to-tail amidation, disulfide bond formation, thioether formation,<sup>40</sup> side-chain-to-tail amidation/N-alkylation, copper (I)-catalyzed azide-alkyne cycloaddition (CuAAC),<sup>41–44</sup> and ring-closing metathesis<sup>45–47</sup> of unnatural, alkenyl side-chains in an  $i$ ,  $i+4$  or  $i$ ,  $i+7$  relationship.<sup>48</sup> Peptides can be prepared either via expression in transformed cells, through the use of purified ribosomes<sup>49</sup> or synthetically through solution phase or on-bead solid phase peptide synthesis (SPPS).<sup>50–52</sup> The heterogeneous synthesis environment of SPPS enables the high-throughput, automated synthesis of high-purity peptides that can contain amino acids not present in the original chiral pool.

Synthetic libraries of cyclic peptides can be prepared through a variety of different methods. Libraries prepared *in vivo* involve transforming cells to generate the library fused to some sort of biomolecule such as phages,<sup>53</sup> ribosomes,<sup>54</sup> mRNA,<sup>55</sup> and DNA.<sup>56</sup> Phage-displayed libraries typically undergo an iterative method of screening called biopanning.<sup>57–59</sup> In this recursive method phages that bind to the desired target are used as the starting point for successively smaller, focused libraries. These libraries are screened

under increasingly more selective conditions until peptides with high binding affinity for the target protein are isolated. After a library screen the plasmids corresponding to library hits are sequenced and the peptides are scaled up for further characterization. For peptides that are prepared in a purely synthetic fashion, the best method for generating a large combinatorial one-bead-one-compound library is the split-and-mix method independently discovered by researchers Furka<sup>60</sup> and Lam<sup>61</sup> (Figure 1.3).

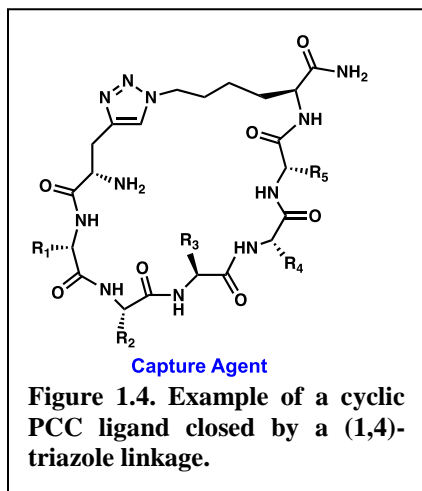


In all peptide libraries the theoretical sequence diversity is roughly  $X^n$ , where  $X$  is the number of amino acids in the residue pool and  $n$  is the length of the peptide chain. As a result peptides six residues long can result in libraries that contain  $>10^7$  unique sequences with peptide libraries of 20mers reaching theoretical sizes of  $>10^{25}$ , which is an incredibly large sequence space. Library screens can occur on-bead with the resulting peptide hits sequenced using Edman degradation<sup>62</sup> on-bead or the peptides are linearized

and cleaved from bead for mass spectrometry analysis methods.<sup>63</sup> The known sequences are then synthesized using standard SPPS protocols<sup>64</sup> and characterized further.

Cyclic peptides have been developed against a variety of protein targets for both *in vivo* and *in vitro* applications.<sup>54,59,65–69</sup>

## Section 1.4-Protein-Catalyzed Capture Agents

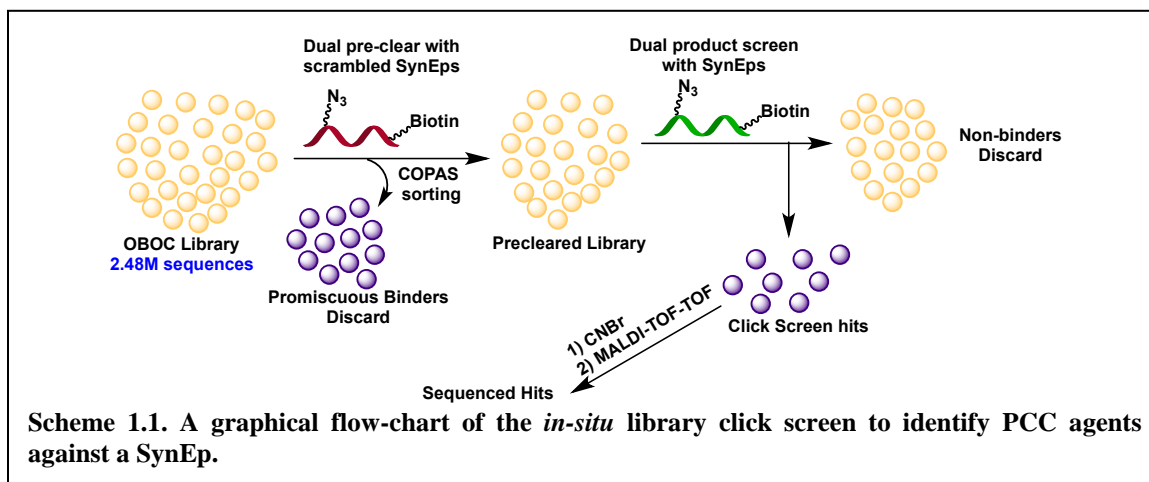


Protein-catalyzed capture (PCC) agents are a specialized class of peptides that were first disclosed by Heath and coworkers in 2009. PCC agents differ from other cyclic peptides in that the library screening process involves a protein/oligopeptide-catalyzed reaction. The PCC technology initially was composed of purely linear ligands, but now consists of cyclic

ligands as the cyclic PCCs exhibit greater stability and binding affinity at the 1° ligand stage (Figure 1.4). The cyclic PCC agents have a variable 5-mer recognition region that is held in place by either a 1,4-triazole, resulting from a copper catalyzed (1,3)-dipolar cycloaddition, or an alkenyl ring closure, resulting from a ring-closing metathesis (RCM).

A PCC library is prepared using SPPS protocols on Tentagel resin using a pool of roughly 18-20 amino acids in a split-and-mix protocol using a pool of roughly 18-20 amino acids to yield a theoretical library size of 1.78 M-3.20 M sequences. Following ring-closure an amino acid containing a click handle is added to the N-terminus amino group and all acid-labile side chain protecting groups are removed.<sup>70</sup> While either an azido-presenting <sup>4-N<sub>3</sub></sup>K (Az4) or alkyne-presenting <sup>α-propargyl</sup>Gly (Pra) residue can be added, the Az4 group is prone to reduction over time to form Lys plus nitrogen gas, and

consequently Pra-presenting libraries are preferred. Screening a library against a particular target protein involves first selecting a solvent-exposed region of the protein that corresponds to a unique signature for the protein e.g. a point mutation for oncogenic



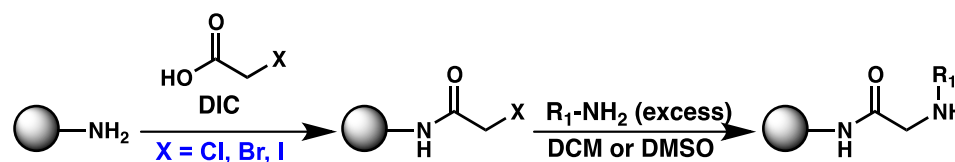
proteins or sequence divergence for polymorphic proteins. This epitope of 9-30 amino acids gets synthesized with both a biotin handle for detection and with one of its amino acids substituted for a structurally similar complementary click handle to prepare a synthetic epitope (SynEp) for screening.<sup>70</sup> Common substitutions are: I, K, L → Az<sup>4</sup> and Pro → <sup>4</sup>-N<sup>3</sup>Pro or G → Pra.

The *in-situ* PCC library screening process involves two separate screening rounds in order to minimize the amount of “sticky” sequences isolated (Scheme 1.1). The *in-situ* click screen utilized is a very low-yielding reaction, which results in the routine isolation of cyclic PCCs with low nM binding affinity for the target protein.<sup>71</sup> First, the library undergoes a pre-clear typically with a scrambled SynEp to identify beads that bind non-selectively. The resulting colored beads are removed either by hand or through an automated complex object parametric analysis and sorting (COPAS) protocol,<sup>72</sup> and the remaining library is incubated with the SynEp to identify hits. The latest version of this screening protocol typically yields 5-10 colored beads after the product screen, which are

then sequenced either using Edman degradation<sup>62</sup> or the more accurate matrix-assisted laser desorption ionization time-of-flight time-of-flight (MALDI-TOF-TOF).<sup>73</sup> Hit sequences are scaled up using standard Fmoc/<sup>t</sup>Bu SPPS protocols<sup>74</sup> with a conjugated biotin assay handle before testing the hits against the full-protein. This evaluation typically involves measuring the binding affinity of the ligands through an immunoassay either through a bulk 96-well plate setup or on the recently disclosed barcoded rapid-assay platform.<sup>75</sup> The best ligands identified by the best binding affinity ( $EC_{50}$  value) are subjected to further characterization/medicinal chemistry optimization in order to arrive at fully optimized ligands for the desired application. Biligands can be formed by raising PCC ligands against multiple regions of the same protein and linking the best ligands through a flexible PEG<sub>n</sub> linker.

Protein catalyzed capture agents have been utilized for both the detection of proteins that function as disease biomarkers<sup>70,76–79</sup> and for exerting a functional effect on a particular protein.<sup>17,75,80–86</sup> A few PCCs have exerted this effect in cellular assays, which suggests that PCC agents could be used *in vivo* as therapeutics particularly as cancer drugs.

## Section 1.5-Peptoids



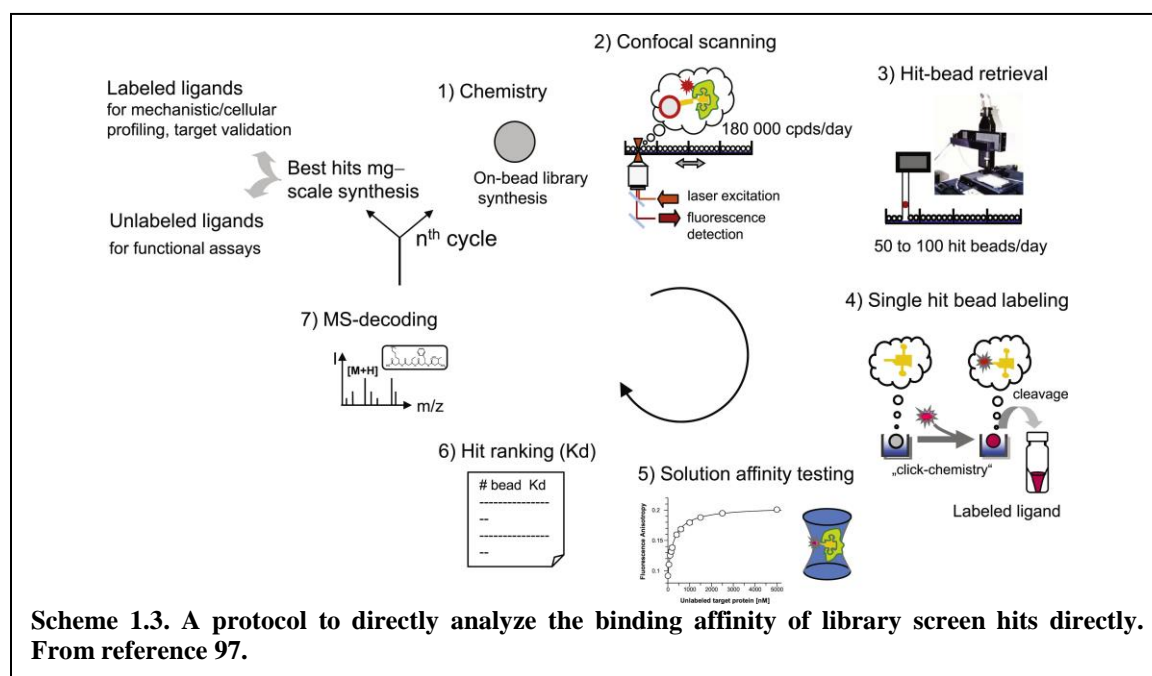
**Scheme 1.2. The submonomer synthesis method for the preparation of peptoids.** The submonomer pathway begins by coupling a 2-halo acid, typically 2-bromoacetic acid through a DIC active-ester intermediate. Next, a nucleophilic  $S_N2$  displacement with an excess of a primary amine installs the N-alkyl group ready for another round of coupling.



Peptoids, or N-substituted polyglycine polymers, are a class of synthetic polymer that was first disclosed in 1992.<sup>18,87</sup> While similar to peptides in terms of having a polyamide backbone, the side chains migrated from the  $\alpha$ -carbon to the amide nitrogen. As a result, the biomimetic polymer has an achiral backbone with tertiary amides that are completely resistant to proteolytic degradation. Peptoids are prepared via a modified solid-phase synthesis method called the submonomer method, which has been adapted to automated synthesis (Scheme 1.2).<sup>88</sup> This iterative two-step process involves acetylation of the free amine group with a 2-halo acetic acid, typically 2-bromo acetic acid, followed by  $S_N2$  displacement of the halogen with a primary amine. The use of a primary amine to introduce functionality greatly expands the pool of potential side chains beyond the original twenty side chains found in  $\alpha$ -amino acids. As a result, non-canonical side chains can be readily introduced. Most early reports of peptoids were linear, but recently cyclization methods have been developed for both solution phase and on-bead cyclization<sup>89</sup> including head-to-tail amide formation,<sup>90-93</sup> side-chain-tail cyclization,<sup>94</sup> side-chain click reactions,<sup>95,96</sup> side-chain ring-closing metathesis (RCM),<sup>47,97</sup> and triazine thioether formation.<sup>98,99</sup> Peptoid libraries are prepared using the aforementioned split-and-mix protocol to prepare OBOC libraries ranging in size from modest ( $\sim 10^{3-4}$ ) to large ( $\sim 10^{5-6}$ ) as the pool of viable amines increases.

Peptoid libraries can be screened by a variety of processes. For example, a classic OBOC library screen protocol for peptoid libraries was reported by Kodadek *et al.*<sup>100</sup> The main difference between this screening process and related ones reported for peptide libraries is the rescreening of hit peptoid beads to distinguish true hits from sticky false positives. Kodadek and coworkers also reported on a two-color cell-based procedure for

identifying peptoids that bind to membrane proteins.<sup>101</sup> In this screening process hits that bind cells with the overexpressed protein, but not the control cells, were taken as hits. Another screening protocol incubates the peptoid library with a His<sub>6</sub>-tagged target protein and captures hits with magnetic anti-His<sub>6</sub> beads. The hits from a library screen can be sequenced in a manner similar to peptides and PCCS: Edman degradation on-bead or tandem mass-spectrometry after cyanogen bromide cleavage from bead.<sup>102</sup> In the case of cyclic peptoids a methionine residue is be incorporated both into the macrocycle and

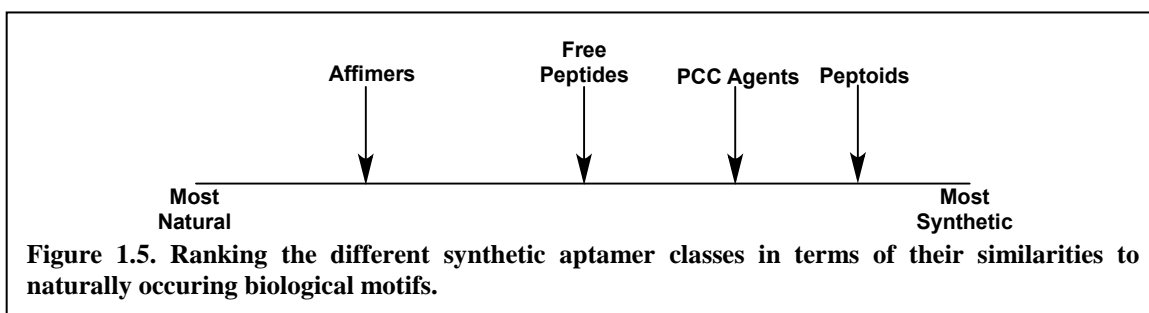


adjacent to the resin linker in order to linearize the hits and cleave from the bead in a single step.<sup>100</sup> Similar to the aforementioned aptamer classes, validation of the scaled-up hits involves measuring their desired performance off-bead and can occur after spotting the hits onto a spin-coated glass slide in a microarray format. However, Weidemann *et al.* disclosed a strategy where the library hits can be analyzed for binding affinity *before* scale-up occurs (Scheme 1.3).<sup>103</sup> After resin cleavage a fluorescent tag gets clicked onto the peptoid hit for measurement of the binding affinity with fluorescence polarization

(FP) measurement. As a result only the most promising hits are subject to scale-up and further investigation.

Peptoids have been raised against a variety of proteins both for detection of proteins in a microarray<sup>104</sup> and for functionally affecting the target protein in a manner that makes them potential therapeutics.<sup>101,105–109</sup> While the above examples highlight the ability of peptoids to target a wide variety of proteins, the affinity is modest likely due to the linear nature of the peptoids. Cyclic peptoids would be expected to have enhanced affinity/functional effects as a result of their more constrained nature.

## Section 1.6-Analysis of the Synthetic Aptamer Classes and Future Outlook



The different aptamer classes can be ranked in terms of their “naturalness” (Figure 1.5). Affimers are the most natural as they contain a protein scaffold similar to antibodies and are typically prepared using cell-based expression systems. Peptides are a transitional aptamer sub-class as peptides can be inspired by naturally occurring peptide sequences and generated via cellular expression, but a growing number of reported peptides are the result of random sequence generation and synthetic SPPS generation. PCCs are a more synthetic sub-class of aptamers relative to free-peptides as the stapling methods used are purely synthetic-based and the ligands are derived solely from combinatorial libraries. However, PCC agents still utilize the chiral pool of amino acids as a starting basis set

even if unnatural amino acids are included. Peptoids represent a purely synthetic class that cannot be traced back to any naturally occurring biomolecules or biopolymers, and a method of synthesis that is purely synthetic although there are limited reports of ribosomes preparing peptoids. While the synthetic aptamer classes are easier to prepare than antibodies, as they do not require the use of animals, only the more synthetic classes are able to incorporate unnatural motifs or easily modify the aptamer through modular substitution of “residues” (Table 1.2).

**Table 1.2. Comparison of structure and preparation of antibodies and synthetic aptamers.**

Property	Peptamers	Peptides	PCC Agents	Peptoids	Antibodies
Typical Size	~15 kDa	0.5-3 kDa	1-3 kDa	0.5-3 kDa	>150 kDa
Preparation	Expression	SPS/Expression	SPS	SPS	Expression
Modular Synthesis?	No	Yes	Yes	Yes	No
Introduction of Unnatural Motifs	Difficult	Easy for SPS preparation	Easy	Very Easy	Difficult
Scalable Synthesis	No	Yes if SPS	Yes	Yes	No
Multimers?	No	Yes	Yes	Yes	Yes

In fact, the automated synthesis methods available to prepare peptides, PCCs, and peptoids, respectively means that is easy to generate the large number of derivative sequences needed to interrogate the structure-activity relationship (SAR) between the recognition regions and desired functionality in a medicinal chemistry optimization approach.

Each aptamer class has physical properties superior to antibodies (Table 1.3) with peptamers and free-peptides the best characterized. The lack of information on immunogenicity for PCC agents and peptoids likely results from their limited use *in vivo*.

The aptamers wherein screening can occur on-bead have screening and sequencing procedures that enable rapid identification of hits without intensive plasmid isolation and screening. However, the use of cell-based screening assays with affimers yields information about functional performance in a complex environment.

**Table 1.3. Comparison of physical properties of antibodies and synthetic aptamers.**

Property	Peptamers	Peptides	PCC Agents	Peptoids	Antibodies
Proteolytic Stability	Low	Medium	Medium	High	Low
Temperature Stability	High	Medium	High	High	Low
Immunogenic?	Unknown	Medium	Unknown	Unknown	High
Cross reactivity	Low	Medium	Low	Unknown	Medium
Cellular Penetration	Yes	Yes	Yes	Yes	No
Nuclear Penetration	Yes	Yes	Unknown	Unknown	No
Typical EC <sub>50</sub> /K <sub>D</sub> values	nM	μM→nM	nM	μM→nM	nM→pM

At first glance the peptoid aptamer subclass appears to be the best suited for *in vivo* applications. In fact, the backbone of peptoids lends superior proteolytic stability as the backbone contains 3° amides. Additionally, the hydrophobic, floppy backbone mimics the structure of small molecules and consequently nearly all peptoids are cellular-penetrant although cyclic peptoids outperform linear peptoids. However, the affinity of peptoids hovers in the low μM to high nM range. The weaker affinity levels can stem from several causes. First, the polyglycine backbone is much more conformationally flexible and binding likely involves a larger entropic penalty than required with the more constrained peptide backbone. Conformational rigidity can be built into peptoids by cyclizing, incorporating bulky, chiral side chains on the amines, and substituting the 2-bromo acetic acid precursor for chiral 2-bromo propionic acids or rigid halogenated N-

heterocyclic acids. However, with the size of peptoid libraries typically in the  $10^4$  to  $10^6$  range, varying the composition of the backbone sharply limits the number of amines can be used to generate a library. Increasing the size of peptoid libraries further would dramatically increase the number of false positives/hits generated in a screen, which would require more resources to validate and identify lead compounds for further characterization. Additionally, most peptoids reported are linear peptoids, which are even more “floppy”, with the number of cyclic peptoid disclosures expected to increase in the near future as most peptoid cyclization methods were reported in the last five years. Both peptides and PCCs have been regularly reported with nM affinity, and they can be engineered to penetrate cells either by making their backbones more rigid/hydrophobic with N-methyl amides or by attaching a cell-penetrating peptide sequence (CPP). However, the library screen of PCC agents generates the fewest hits, and, when Edman degradation artifacts are ignored, the lowest number of false positives of the discussed aptamer subclasses. In terms of overall stability and functionality PCC agents represent a “Goldilocks” solution between free-peptides and peptoids for *in vivo* applications.

For microarray development, while each aptamer class can be immobilized in an array format, the affimer class has a growing body of literature on their use as aptasensors, which makes affimers attractive for the development of proteomic microarrays. However, biotinylated PCC agents can be complexed with streptavidin-DNA conjugates for *indirect* assembly onto a microarray, which means that PCC agents can be independently prepared and modularly immobilized to the platform right before running an assay. This rapid assembly of separately generated components was demonstrated with the B-RAP technology and could be adapted to prepare highly modular PCC-based proteomic

analysis kits that can be rapidly adapted to the needs of the particular proteomic panel. To our knowledge the other aptamer classes are directly assembled onto the microarray platform, which requires the preparation of an entirely new microarray if a new protein panel is desired. As a result, the use of PCC agents in diagnostic microarrays and the advancement of PCCs into therapeutic clinical trials is expected to increase dramatically in the next 10-15 years.

## Section 1.7-The Theme of the Thesis

The work summarized in the remainder of this thesis represents advances to the PCC technology. Chapter 2 summarizes contributions to an ongoing effort to develop a high-throughput production pipeline for PCC ligands with the development of a barcoded rapid assay platform (B-RAP) technology for the simultaneous evaluation of the binding affinity of up to fifteen different PCC agents in a single day. This chapter also discloses progress towards developing PCC agents against *difficult* proteins with the identification of PCC agents that bind to and inhibit the enzymatic activity of Kirsten rat sarcoma (KRas) protein, whose oncogenic variants drive 20-25% of all cancers, but lack FDA-approved drug treatments. Chapter 3 summarizes another effort to target *challenging* protein targets with the development of PCC agents against the sticky *unstructured*, highly polymorphic protein Histidine rich protein II (HRP2).

## Section 1.8-References

- (1) Williams, C. M. M.; Galli, S. J. *J. Allergy Clin. Immunol.* **2000**, *105* (5), 847.
- (2) Diamantis, N.; Banerji, U. *Br. J. Cancer* **2016**, *114* (4), 362.
- (3) Melero, I.; Berman, D. M.; Aznar, M. A.; Korman, A. J.; Gracia, J. L. P.; Haanen, J. *Nat. Rev. Cancer* **2015**, *15* (8), 457.

- (4) Shih, K.; Arkenau, H.-T.; Infante, J. R. *Drugs* **2014**, *74* (17), 1993.
- (5) Shin, Y. S.; Ahmad, H.; Shi, Q.; Kim, H.; Pascal, T. A.; Fan, R.; Goddard, W. A.; Heath, J. R. *ChemPhysChem* **2010**, *11* (14), 3063.
- (6) Bailey, R. C.; Kwong, G. A.; Radu, C. G.; Witte, O. N.; Heath, J. R. *J. Am. Chem. Soc.* **2007**, *129* (7), 1959.
- (7) Fan, R.; Vermesh, O.; Srivastava, A.; Yen, B. K. H.; Qin, L.; Ahmad, H.; Kwong, G. A.; Liu, C. C.; Gould, J.; Hood, L.; Heath, J. R. *Nat. Biotechnol.* **2008**, *26* (12), 1373.
- (8) Travers, P.; Walport, M.; Schlomchik, M.; Janeway, C. In *Immunobiology: The Immune System in Health and Disease*; Garland Science: New York, 2001.
- (9) Gossow, D.; Seemann, G. *Contrib. to Ontol. Blood Nucl. Med. Commun. Trends Biotechnol. Proc. Natl. Acad. Sci. U.S.A. Adv. Immunol. Nat.* **1988**, *8* (643), 227.
- (10) Chadd, H. E.; Chamow, S. M. *Curr. Opin. Biotechnol.* **2001**, *12* (2), 188.
- (11) Holliger, P.; Hudson, P. J. *Nature Biotechnology*. 2005, pp 1126–1136.
- (12) Eyer, L.; Hruska, K.; Eyer, L.; Hruska, K. *Vet. Med. (Praha)*. **2012**, *57* (9), 439.
- (13) Ellington, A. D.; Szostak, J. W. *Nature* **1990**, *346*, 818.
- (14) Keefe, A. D.; Pai, S.; Ellington, A. *Nat. Rev. Drug Discov.* **2010**, *9* (7), 537.
- (15) Fosgerau, K.; Hoffmann, T. *Drug Discov. Today* **2015**, *20* (1), 122.
- (16) Colas, P. *J. Biol.* **2008**, *7* (1), 2.
- (17) Agnew, H. D.; Rohde, R. D.; Millward, S. W.; Nag, A.; Yeo, W. S.; Hein, J. E.; Pitram, S. M.; Abdul Ahad Tariq, V.; Burns, A. M.; Krom, R. J.; Fokin, V. V.; Barry Sharpless, K.; Heath, J. R. *Angew. Chemie Int. Ed.* **2009**, *48* (27), 4944.
- (18) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.;



- Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K.; Spellmeyer, D. C.; Tans, R.; Frankel, A. D.; Santi, D. V.; Cohen, F. E.; Bartlett, P. A. *Chemistry (Easton)*. **1992**, 89, 9367.
- (19) Colas, P.; Cohen, B.; Jessen, T.; Grishina, I.; McCoy, J.; Brent, R. *Nature* **1996**, 380 (6574), 548.
- (20) Woodman, R.; Yeh, J. T.-H. H.; Laurenson, S.; Ko Ferrigno, P. *J. Mol. Biol.* **2005**, 352 (5), 1118.
- (21) Peelle, B.; Gururaja, T. L.; Payan, D. G.; Anderson, D. C. *J. Protein Chem.* **2001**, 20 (6), 507.
- (22) Abedi, M. R.; Caponigro, G.; Kamb, A. *Nucleic Acids Res.* **1998**, 26 (2), 623.
- (23) Beste, G.; Schmidt, F. S.; Stibora, T.; Skerra, A. *Biochemistry* **1999**, 96, 1898.
- (24) Tiede, C.; Tang, A. A. S.; Deacon, S. E.; Mandal, U.; Nettleship, J. E.; Owen, R. L.; George, S. E.; Harrison, D. J.; Owens, R. J.; Tomlinson, D. C.; McPherson, M. *J. Protein Eng. Des. Sel.* **2014**, 27 (5), 145.
- (25) Hoffmann, T.; Stadler, L. K. J.; Busby, M.; Song, Q.; Buxton, A. T.; Wagner, S. D.; Davis, J. J.; Ko Ferrigno, P. *Protein Eng. Des. Sel.* **2010**, 23 (5), 403.
- (26) Borghouts, C.; Kunz, C.; Groner, B.; Borghouts, C.; Kunz, C. *Comb. Chem. High Throughput Screen.* **2008**, 11 (2), 135.
- (27) Field, S.; Song, O. *Nature* **1989**, 340, 245.
- (28) Auerbach, D.; Stagljar, I.; Auerbach, D.; Stagljar, I. In *Proteomics and Protein-Protein Interactions: Biology, Chemistry, Bioinformatics, and Drug Design*; Springer, 2005; pp 19–31.
- (29) Tavernier, J.; Eyckerman, S.; Lemmens, I.; Van Der Heyden, J.; Vandekerckhove,

- J.; Van Ostade, X. *Clin. Exp. Allergy* **2002**, 32 (10), 1397.
- (30) Lemmens, I.; Lievens, S.; Tavernier, J. *Biochem. Soc. Trans.* **2008**, 36 (6).
- (31) Genomics, B.; Fiebitz, A.; Nyarsik, L.; Haendler, B.; Hu, Y.-H.; Wagner, F.; Thamm, S.; Lehrach, H.; Janitz, M.; Vanhecke, D. *BMC Genomics* **2008**, 9 (9).
- (32) Vazquez-Lombardi, R.; Phan, T. G.; Zimmermann, C.; Lowe, D.; Jermutus, L.; Christ, D. *Drug Discovery Today*. 2015, pp 1271–1283.
- (33) Evans, D.; Johnson, S.; Laurenson, S.; Davies, G.; Ko Ferrigno, P.; Walti, C. *J. Biol.* **2008**, 7 (1), 3.
- (34) Estrela, P.; Paul, D.; Li, P.; Keighley, S. D.; Migliorato, P.; Laurenson, S.; Ferrigno, P. K. *Electrochim. Acta* **2008**, 53 (22), 6489.
- (35) Zhuravski, P.; Arya, S. K.; Jolly, P.; Tiede, C.; Tomlinson, D. C.; Ko Ferrigno, P.; Estrela, P. *Biosens. Bioelectron.* **2018**, 108, 1.
- (36) Sharma, R.; Deacon, S. E.; Nowak, D.; George, S. E.; Szymonik, M. P.; Tang, A. A. S.; Tomlinson, D. C.; Davies, A. G.; McPherson, M. J.; Wälti, C. *Biosens. Bioelectron.* **2016**, 80, 607.
- (37) Weckman, N. E.; McRae, C.; Ko Ferrigno, P.; Seshia, A. A. *Analyst* **2016**, 141 (22), 6278.
- (38) Johnson, A.; Song, Q.; Ko Ferrigno, P.; Bueno, P. R.; Davis, J. J. *Anal. Chem.* **2012**, 84 (15), 6553.
- (39) Berto, M.; Diacci, C.; D'Agata, R.; Pinti, M.; Bianchini, E.; Lauro, M. Di; Casalini, S.; Cossarizza, A.; Berggren, M.; Simon, D.; Spoto, G.; Biscarini, F.; Bortolotti, C. A. *Adv. Biosyst.* **2017**, 2, 1700072.
- (40) Smeenk, L. E. J.; Dailly, N.; Hiemstra, H.; Van Maarseveen, J. H.; Timmerman, P.

- Org. Lett.* **2012**, *14* (5), 1194.
- (41) Lau, Y. H.; Wu, Y.; Rossmann, M.; Tan, B. X.; De Andrade, P.; Tan, Y. S.; Verma, C.; McKenzie, G. J.; Venkitaraman, A. R.; Hyvönen, M.; Spring, D. R. *Angew. Chemie - Int. Ed.* **2015**, *54* (51), 15410.
- (42) Punna, S.; Kuzelka, J.; Wang, Q.; Finn, M. G. *Angew. Chemie - Int. Ed.* **2005**, *44* (15), 2215.
- (43) Li, H.; Aneja, R.; Chaiken, I. *Molecules*. 2013, pp 9797–9817.
- (44) Ingale, S.; Dawson, P. E. *Org. Lett.* **2011**, *13* (11), 2822.
- (45) Kim, Y.; Grossmann, T. N.; Verdine, G. L. *Nat. Protoc.* **2011**, *6* (6), 761.
- (46) Stringer, J. R.; Crapster, J. A.; Guzei, I. A.; Blackwell, H. E. *J. Org. Chem. J. Org. Chem* **2010**, *75* (18), 6068.
- (47) Khan, S. N.; Kim, A.; Grubbs, R. H.; Kwon, Y.-U. *Org. Lett.* **2011**, *13* (7), 1582.
- (48) Blackwell, H. E.; Sadowsky, J. D.; Howard, R. J.; Sampson, J. N.; Chao, J. A.; Steinmetz, W. E.; O’Leary, D. J.; Grubbs, R. H. *J. Org. Chem.* **2001**, *66* (16), 5291.
- (49) Bashiruddin, N. K.; Suga, H. *Curr. Opin. Chem. Biol.* **2015**, *24*, 131.
- (50) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85* (14), 2149.
- (51) Merrifield, R. B. *Angew. Chemie (International ed English)* **1985**, *24* (10), 799.
- (52) Carpino, L. A.; Han, G. Y. *J. Am. Chem. Soc.* **1970**, *92* (19), 5748.
- (53) Morioka, T.; Loik, N. D.; Hipolito, C. J.; Goto, Y.; Suga, H. *Curr. Opin. Chem. Biol.* **2015**, *26*, 34.
- (54) Wada, A. *Front. Immunol.* **2013**, *4* (AUG), 1.
- (55) Cotten, S. W.; Zou, J.; Valencia, C.; Liu, R. *Nat. Protoc.* **2011**, *6* (8), 1163.
- (56) Odegrip, R.; Coomber, D.; Eldridge, B.; Hederer, R.; Kuhlman, P. A.; Ullman, C.;

- FitzGerald, K.; McGregor, D. *Proc. Natl. Acad. Sci.* **2004**, *101* (9), 2806.
- (57) Parmley, S. F.; Smith, G. P. *Gene* **1988**, *73* (2), 305.
- (58) Mandecki, W.; Chen, Y.-C. C.; Grihalde, N. *J Theor Biol* **1995**, *176* (4), 523.
- (59) Deyle, K.; Kong, X.-D.; Heinis, C. *Acc. Chem. Res.* **2017**, *50* (8), 1866.
- (60) Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Pept. Protein Res.* **1991**, *37* (6), 487.
- (61) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M. > M.; Knapp, R. J.; Herah, E. M.; Heruby, V. J.; Kazmierski, W. M. > M.; Knapp, R. J. *Nature* **1991**, *354* (6348), 82.
- (62) Edman, P.; Begg, G. *Eur. J. Biochem.* **1967**, *1* (1), 80.
- (63) Hernandez, P.; Müller, M.; Appel, R. D. *Mass Spectrom. Rev.* **2006**, *25* (2), 235.
- (64) Palomo, J. M. *RSC Adv.* **2014**, *4* (62), 32658.
- (65) Walport, L. J.; Obexer, R.; Suga, H. *Curr. Opin. Biotechnol.* **2017**, *48*, 242.
- (66) Lennard, K. R.; Tavassoli, A. *Chem. - A Eur. J.* **2014**, *20* (34), 10608.
- (67) Trinh, T. B.; Upadhyaya, P.; Qian, Z.; Pei, D. *ACS Comb. Sci.* **2016**, *18* (1), 75.
- (68) Osher, E. L.; Castillo, F.; Elumalai, N.; Waring, M. J.; Pairaudeau, G.; Tavassoli, A. *Bioorganic and Medicinal Chemistry*. March 2018,.
- (69) Qian, Z.; Dougherty, P. G.; Pei, D. *Curr. Opin. Chem. Biol.* **2017**, *38* (38), 80.
- (70) Das, S.; Nag, A.; Liang, J.; Bunck, D. N.; Umeda, A.; Farrow, B.; Coppock, M. B.; Sarkes, D. A.; Finch, A. S.; Agnew, H. D.; Pitram, S.; Lai, B.; Yu, M. B.; Museth, A. K.; Deyle, K. M.; Lepe, B.; Rodriguez-Rivera, F. P.; McCarthy, A.; Alvarez-Villalonga, B.; Chen, A.; Heath, J.; Stratis-Cullum, D. N.; Heath, J. R. *Angew. Chemie Int. Ed.* **2015**, *54* (45), 13219.

- (71) Manetsch, R.; Krasiński, A.; Radić, Z.; Raushel, J.; Taylor, P.; Sharpless, K. B.; Kolb, H. C. *J. Am. Chem. Soc.* **2004**, *126* (40), 12809.
- (72) Bunck, D. N.; Heath, J. R. *Unpublished*.
- (73) Lee, S. S.; Lim, J.; Tan, S.; Cha, J.; Yeo, S. Y.; Agnew, H. D.; Heath, J. R. *Anal. Chem.* **2010**, *82* (2), 672.
- (74) Atherton, E.; Fox, H.; Harkiss, D.; Logan, C. J.; Sheppard, R. C.; Williams, B. J. *J. Chem. Soc. Chem. Commun.* **1978**, No. 13, 537.
- (75) McCarthy, A. M.; Kim, J.; Museth, A. K.; Henning, R. K.; Heath, J. E.; Winson, E.; Oh, J. J.; Heath, J. R. *An Allosteric Inhibitor of KRas Identified Using a Barcoded Rapid Assay Microchip Platform*; 2018.
- (76) Pfeilsticker, J. A.; Umeda, A.; Farrow, B.; Hsueh, C. L.; Deyle, K. M.; Kim, J. T.; Lai, B. T.; Heath, J. R. *PLoS One* **2013**, *8* (10), 4.
- (77) Farrow, B.; Hong, S. A.; Romero, E. C.; Lai, B.; Coppock, M. B.; Deyle, K. M.; Finch, A. S.; Stratis-Cullum, D. N.; Agnew, H. D.; Yang, S.; Heath, J. R. *ACS Nano* **2013**, *7* (10), 9452.
- (78) Lai, B. T.; Wilson, J. A.; Malette Lored, J.; Pitram, S. M.; LaBerge, N. A.; Heath, J. R.; Agnew, H. D. *Chem. - A Eur. J.* **2018**, *24* (15), 3760.
- (79) Coppock, M. B.; Warner, C. R.; Dorsey, B.; Orlicki, J. A.; Sarkes, D. A.; Lai, B. T.; Pitram, S. M.; Rohde, R. D.; Malette, J.; Wilson, J. A.; Kearney, P.; Fang, K. C.; Law, S. M.; Candelario, S. L.; Farrow, B.; Finch, A. S.; Agnew, H. D.; Heath, J. R.; Stratis-Cullum, D. N. *Biopolymers* **2017**, *108* (2), 1.
- (80) Bunck, D. N.; Atsavapranee, B.; Museth, A. K.; Vandervelde, D.; Heath, J. R.; Bunck, D. N.; Heath, J. R. *Angew. Chem. Int. Ed. Angew. Chem* **2018**, *10* (10).

- (81) Liang, J.; Bunck, D. N.; Mishra, A.; Idso, M. N.; Heath, J. R. *Inhibition of heme sequestration of Histidine-Rich Protein 2 using multiple epitope-targeted peptides*; 2018.
- (82) Farrow, B.; Wong, M.; Malette, J.; Lai, B.; Deyle, K. M.; Das, S.; Nag, A.; Agnew, H. D.; Heath, J. R. *Angew. Chemie Int. Ed.* **2015**, *54* (24), 7114.
- (83) Millward, S. W.; Henning, R. K.; Kwong, G. A.; Pitram, S.; Agnew, H. D.; Deyle, K. M.; Nag, A.; Hein, J.; Lee, S. S.; Lim, J.; Pfeilsticker, J. A.; Sharpless, K. B.; Heath, J. R. *J. Am. Chem. Soc.* **2011**, *133* (45), 18280.
- (84) Nag, A.; Das, S.; Liu, F.; Kim, J.; Goddard, W. A.; Heath, J. R. *Submitted* **2018**.
- (85) Deyle, K. M.; Farrow, B.; Hee, Y. Q.; Work, J.; Wong, M.; Lai, B.; Umeda, A.; Millward, S. W.; Nag, A.; Das, S.; Heath, J. R.; Qiao Hee, Y.; Work, J.; Wong, M.; Lai, B.; Umeda, A.; Millward, S. W.; Nag, A.; Das, S.; Heath, J. R. *Nat. Chem.* **2015**, *7* (5), 455.
- (86) Henning, R. K.; Varghese, J. O.; Das, S.; Nag, A.; Tang, G.; Tang, K.; Sutherland, A. M.; Heath, J. R. *J. Pept. Sci.* **2016**, *22* (4), 196.
- (87) Zuckermann, R. N. *Biopolymers* **2011**, *96* (5), 545.
- (88) Zuckermann, R. N.; Kerr, J. M.; Moosf, W. H.; Kent, S. B. H. *Journal of the American Chemical Society*. 1992, pp 10646–10647.
- (89) Webster, A. M.; Cobb, S. L. *Chem. - A Eur. J.* **2018**, *24*, Accepted.
- (90) Shin, S. B. Y.; Yoo, B.; Todaro, L. J.; Kirshenbaum, K. *J. Am. Chem. Soc.* **2007**, *129* (11), 3218.
- (91) D'Amato, A.; Volpe, R.; Vaccaro, M. C.; Terracciano, S.; Bruno, I.; Tosolini, M.; Tedesco, C.; Pierri, G.; Tecilla, P.; Costabile, C.; Della Sala, G.; Izzo, I.; De

- Riccardis, F. *J. Org. Chem.* **2017**, 82 (17), 8848.
- (92) Andreev, K.; Martynowycz, M. W.; Ivankin, A.; Huang, M. L.; Kuzmenko, I.; Meron, M.; Lin, B.; Kirshenbaum, K.; Gidalevitz, D. *Langmuir* **2016**, 32 (48), 12905.
- (93) Culf, A. S.; Čuperlović-Culf, M.; Léger, D. A.; Decken, A. *Org. Lett.* **2014**, 16 (10), 2780.
- (94) Kaniraj, P. J.; Maayan, G. *Org. Lett.* **2015**, 17 (9), 2110.
- (95) Salvador, C. E. M.; Pieber, B.; Neu, P. M.; Torvisco, A.; Kleber Z. Andrade, C.; Kappe, C. O. *J. Org. Chem.* **2015**, 80 (9), 4590.
- (96) Chirayil, S.; Luebke, K. J. *Tetrahedron Lett.* **2012**, 53, 726.
- (97) Holub, J. M.; Jang, H.; Kirshenbaum, K. *Org. Lett.* **2007**, 9 (17), 3275.
- (98) Lee, J. H.; Kim, H.-S. S.; Lim, H.-S. S. *Org. Lett.* **2011**, 13 (19), 5012.
- (99) Shin, M. K.; Hyun, Y. J.; Lee, J. H.; Lim, H. S. *ACS Comb. Sci.* **2018**, 20 (4), 237.
- (100) Simpson, L. S.; Kodadek, T. *Tetrahedron Lett.* **2012**, 53 (18), 2341.
- (101) Udugamasooriya, D. G.; Dineen, S. P.; Brekken, R. A.; Kodadek, T. *J. Am. Chem. Soc.* **2008**, 130 (17), 5744.
- (102) Alluri, P. G.; Reddy, M. M.; Bachhawat-Sikder, K.; Olivos, H. J.; Kodadek, T. *J. Am. Chem. Soc.* **2003**, 125 (46), 13995.
- (103) Hintersteiner, M.; Kimmerlin, T.; Kalthoff, F.; Stoeckli, M.; Garavel, G.; Seifert, J.-M.; Meisner, N.-C.; Uhl, V.; Buehler, C.; Weidemann, T.; Auer, M. *Chem. Biol.* **2009**, 16, 724.
- (104) Raveendra, B. L.; Wu, H.; Baccala, R.; Reddy, M. M.; Schilke, J.; Bennett, J. L.; Theofilopoulos, A. N.; Kodadek, T. *Chem. Biol.* **2013**, 20 (3), 351.

- (105) Gao, Y.; Kodadek, T. *ACS Comb. Sci.* **2015**, *17* (3), 190.
- (106) Vendrell-navarro, G.; Rffla, F.; Bujons, J.; Brockmeyer, A. *ChemBioChem* **2015**, *16*, 1580.
- (107) Lim, H. S.; Archer, C. T.; Kodadek, T. *J. Am. Chem. Soc.* **2007**, *129* (25), 7750.
- (108) Simon, R. J.; Kania, R. S.; Zuckermann, R. N.; Huebner, V. D.; Jewell, D. A.; Banville, S.; Ng, S.; Wang, L.; Rosenberg, S.; Marlowe, C. K. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (4), 2374.
- (109) Seo, J.; Lee, B.-C.; Zuckermann, R. N. *Compr. Biomater.* **2011**, *2*, 53.



## Chapter 2

### An Allosteric Inhibitor of KRas Identified Using a Barcoded Rapid Assay Microchip Platform

*"Reproduced in part with permission from Analytical Chemistry, submitted for publication. Unpublished work copyright 2018 American Chemical Society."*

Protein catalyzed capture (PCC) agents are synthetic antibody surrogates that can target a wide variety of biologically relevant proteins. As a step towards developing a high-throughput PCC pipeline we report on the preparation of a barcoded rapid assay platform, used here for the analysis of hits from an *in situ* click screen of a macrocycle peptide library against allosteric epitopes of the KRas protein. The platform utilizes patterned, micrometer scale barcodes composed of orthogonal ssDNA strands on a glass slide. The slide is partitioned into microwells, each of which contains multiple copies of the full barcode. Biotinylated candidate PCCs from the screen are assembled onto designated barcode stripes using a complementary ssDNA-encoded cysteine-modified streptavidin library. A single microchip was utilized for the simultaneous evaluation of fifteen PCC candidate fractions under more than a dozen different assay conditions. The platform permitted a more than a 10-fold savings in time and a more than 100-fold reduction in biological and chemical reagents, relative to traditional multi-well plate assays. The platform was utilized to identify a KRas ligand that exhibits an *in vitro* inhibition constant ( $IC_{50}$ ) of  $\sim 24 \mu M$ , which is an excellent starting point drugging this challenging target.

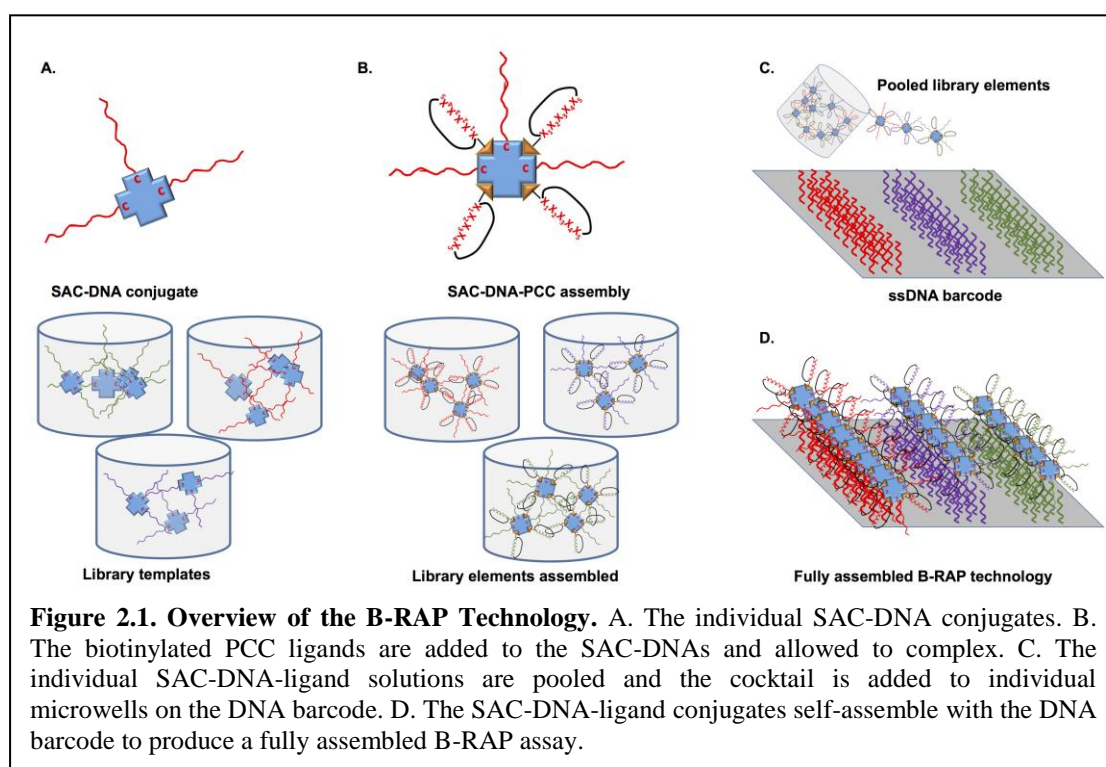
## **An Allosteric Inhibitor of KRas Identified Using a Barcoded Rapid Assay Microchip Platform**

### **Section 2.1-Introduction**

Protein-catalyzed capture agents (PCCs) have been demonstrated to mimic the epitope targeting ability and high avidity of monoclonal antibodies for a number of protein targets.<sup>1</sup> PCCs can be engineered to have combined properties that are difficult to achieve for biologics, such as combinations of physical and biological stability, or, in one example, cell penetration.<sup>2</sup> State-of-the-art PCCs are identified by carrying out an *in situ* click screen<sup>3</sup> of a synthetic, strategically modified polypeptide fragment (the synthetic epitope, or SynEp) of the protein target against a synthetic one-bead-one compound (OBOC) library of macrocyclic peptides. The comprehensive OBOC library typically contains the roughly two million sequences that result from using all combinations of an 18-20 amino acid basis set to construct the variable 5-mer portion of the peptide.

PCC lead compounds are identified through a multi-step process, much of which is highly efficient. The OBOC library is first cleared of non-selective binders by screening against designated interferents. Candidate binders are then identified via a single generation *in situ* click screen against one or more SynEps of the targeted protein. That screen typically yields five to ten hits per SynEp. Once identified, those hit peptides are cleaved from the bead and sequenced using Edman degradation or mass spectrometry, prepared in ~1 mg quantities, and then chromatographically purified. These steps are relatively efficient, and, with commercial robotics, can be accomplished in a few days. However, each PCC candidate must then be tested for binding to the full-length

protein, often in various levels of serum background and under different blocking conditions. These assays are carried out on 96-well plates using a sandwich Enzyme-Linked Immunosorbent Assay (ELISA) format, and represent a limiting factor in PCC production. Consider an *in situ* click screen in which a single OBOC library is screened against two SynEps to yield 15 PCC candidates. Each candidate is tested in, for example, a 10-point binding assay (run in triplicate) against the target protein. This yields  $15 \times 3 \times 10 = 450$  data points, which might be repeated for various background serum concentrations. In addition to being laborious, these assays also consume significant amounts of chemical and biological reagents. Finding a more efficient solution for carrying out such assays should be useful for the production of other artificial antibody-type ligands, such as other classes of



peptides or aptamers.<sup>4-7</sup>

We report here on the barcoded rapid assay platform (B-RAP) (Figure 2.1), which is a microchip platform designed so that an entire set of candidate PCC

ligands may be rapidly evaluated in parallel, using minimal quantities of reagents. Simultaneous testing of all PCCs under identical environments means that all assays are subject to the same uncertainties, which permits ready comparison of the EC<sub>50</sub> values for the entire set of hit peptides. The B-RAP technology draws from the Nucleic Acid Cell Sorting (NACS)<sup>8</sup> and DNA-Encoded Antibody Library (DEAL) methods.<sup>9–12</sup> The B-RAP process starts with a microscope slide that is patterned, using microfluidic flow channels, with a distinct set of orthogonal ssDNA oligomers. The PCC candidates are prepared with a biotin label, and then assembled onto cysteine-modified streptavidin (SAC) scaffolds that have been labelled with complementary ssDNA oligomers.<sup>13–15</sup> Once assembled, these reagents are combined into a cocktail, and assembled onto specific stripes of the barcode pattern using DNA hybridization.<sup>16</sup> The microchip surface itself is partitioned into microliter volume wells, each of which contains multiple copies of the full barcode. The B-RAP technology can simultaneously assay a full panel of candidate PCCs over a range of target protein concentrations (or other conditions), such that the EC<sub>50</sub> binding values for each candidate PCC are concurrently measured.

We used the B-RAP technology to analyze the resulting hits from an epitope targeted *in situ* click screen against the Kirsten rat sarcoma (KRas) protein.<sup>17</sup> Oncoprotein variants of KRas are implicated in driving ~20-25% of all human cancers including almost all pancreatic cancers.<sup>18</sup> Oncogenic Ras proteins have largely evaded targeting by traditional therapeutic techniques,<sup>19–22</sup> but recent work has shown that specific mutant isoforms may be targetable.<sup>23,24</sup> We targeted conserved epitopes denoted Switch I (aa 25-40) and Switch II (aa 56-75), which are known to allosterically influence KRas activity.<sup>25</sup> To our

knowledge, these allosteric regions have not previously been targeted, perhaps because there is no obvious hydrophobic binding pocket. After screening, we tested the resultant hit compounds for their relative binding strengths. The strongest binders were then tested in a functional assay for *in vitro* KRas GTPase activity inhibition.

## Section 2.2-Methods<sup>26</sup>

### 2.2.1-Preparation of the Barcode Rapid Assay Platform

DNA flow-patterned barcode chips, biotinylated peptides, and SAC-DNA were all used to assemble a miniaturized barcode of candidate PCCs for testing in a surface Immunofluorescent assay (IFA). Microfluidic flow patterning of 50  $\mu\text{m}$  wide, 100  $\mu\text{m}$  pitch ssDNA barcodes starts with adhering a polydimethylsiloxane (PDMS) microchannel mold onto a poly-L-lysine (PLL) coated glass microscope slide (Figure A2B.1, Appendix 2B). Reagents were flowed through the microchannels using a “pins-and-tubing-free” system that greatly simplified the preparation of barcoded microchips relative to the previous protocols (Figure A2B.2 and Table A2C.1, Appendix 2B and Appendix 2C).<sup>27,28</sup> The PDMS mold was patterned with microwells at each microchannel inlet (Figure A2B.1A (i, ii), Appendix 2B). Reagents (3-5  $\mu\text{L}$ ) are micropipetted into the wells, and two machined acrylic plates are clamped across the top and bottom of the inlet region. The top acrylic plate contains a cavity that encompasses all of the inlet microwells. This cavity is pressurized to fill the microchannels in about 20 minutes (Figure A2B.2B, Appendix 2B). The increased pressure tolerance of the design can enable the use of microchannels of widths as small as 10  $\mu\text{m}$ . Initially 3 $\mu\text{L}$  of poly-l-lysine (0.1% (<sup>w</sup>/<sub>w</sub>) in H<sub>2</sub>O) is flow patterned and dried overnight before flowing 5 $\mu\text{L}$  of 300  $\mu\text{M}$  of each

ssDNA (Table A2C.2, Appendix 2C) with 2mM bis(sulfosuccinimidyl)suberate (BS3) crosslinker. Approximately twenty to twenty-five DNA barcoded chips may be prepared in parallel (Figure A2B.2C, Appendix 2B). The bottom edge of the barcode is used to validate the coverage density and uniformity of the molecular patterns using fluorophore-labelled complementary ssDNA (Figures A2B.2 and A2B.3, Appendix 2B). Once validated, the barcoded slides may be vacuum-sealed for up to six months storage before use (Figure A2B.4, Appendix 2B; Table A2C.3, Appendix 2C).

The second component of the B-RAP technology, which is also independent of the specific identities of any PCC candidates to be tested, is the library of DNA-bound SAC (SAC-DNA) conjugates used to assemble individual biotinylated PCC candidates onto specific barcode lanes. The SAC protein (see Supplementary Methods in Appendix 2A) was conjugated with ssDNA strands complementary to the barcode DNA oligomers. This was done with N-succinimidyl-4-formylbenzaldehyde (S-4FB) and maleimide 6-hydrazinonicotinamide (MHPH), followed by fast protein liquid chromatography (FPLC) purification (Figure A2B.4, Appendix 2B).

The performance of the library of fifteen SAC-DNAs<sup>16,29</sup> was evaluated by hybridizing library elements onto the flow patterned ssDNA barcodes. The barcodes were then incubated with varying amounts of the fluorophore probe biotin-A<sub>20</sub>-Cy3 (Biotin\*, 50-400nM) (Figure A2B.5, Appendix 2B). The resulting surface fluorescence was measured and compared to the fluorescence signal from the bottom edge barcode validation region. The fluorescent output with 532 nm excitation ( $F_{532}$ ) of the captured biotinylated probe was lower than

that of the validation region (45 to 65k fluorescence units (f.u.)), likely reflecting the size of the SAC protein relative to the Cy3 fluorophore.

### ***2.2.2-KRas Protein Expression and Purification***

The KRas protein isoform 4B was expressed from transformed BL21(D3) *E. coli* cells as a His<sub>6</sub>-tagged protein<sup>30</sup> and purified by FPLC using a Ni-NTA resin (Figure A2B.6, Appendix 2B). The fractions with pure KRas protein were dialyzed into tris-buffered saline (TBS, pH=7.4), aliquotted, and stored at -80.0 °C until needed.

### ***2.2.3-Preparation of Switch I and Switch II SynEps and Scrambled SynEps***

The synthetic epitopes (**SynEp1** and **SynEp2**) were 11-12 amino acid polypeptides with sequences extracted from the allosteric switch regions of KRas (Figure 2.2 and Table A2C.3, Appendix 2C). The **SynEp1** differs from the wild-type sequence as it is missing a valine residue. An azido click handle was added by substituting residue-similar azido-amino acids, as shown in Figure 2.2A. Rearranged versions of the SynEps were also prepared, and used in a pre-screen step to remove promiscuous binders. All epitopes were synthesized on biotin Novatag resin and purified using semi-preparative high performance liquid chromatography (semi-prep HPLC). The appropriate fractions were identified using matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) (Figures A2B.7-A2B.19, Appendix 2B). Each SynEp was dissolved in dimethyl sulfoxide (DMSO), quantified using a Nanodrop 2000 spectrophotometer, and stored at 4 °C until use.

### ***2.2.4-Library Preparation and In-situ Library Click Screen***

A comprehensive OBOC library of 5-mer variable peptide macrocycles, using an 18 amino acid basis set, was prepared as previously reported.<sup>1</sup> The macrocyclic peptides were closed with a 1,4 triazole using Cu(I)-catalyzed click chemistry. These macrocycles were designed to present a propargylglycine click handle. The *in situ* library click screen was a dual SynEp version of a previously reported protocol.<sup>1</sup> After removing the beads that bound to the scrambled SynEps during a pre-clear screen the remaining library was incubated with both **SynEp1** and **SynEp2** (Supplementary Experimental Methods in Appendix 2A). After incubating with an anti-biotin capture antibody and an alkaline-phosphatase conjugated secondary antibody, the hit beads were identified by their deep purple color. The isolated hit beads were stored at RT in 0.1 M hydrochloric acid. Just prior to sequencing by Edman degradation, the beads were decolorized in N-methyl 2-pyrrolidone (NMP) (Table A2C.3, Appendix 2C). The hit compounds were then scaled up on biotin Novatag resin following previously established protocols,<sup>1</sup> purified, lyophilized, reconstituted in DMSO, quantified, and then stored at 4 °C until ready for use.

#### ***2.2.5-Surface Immunofluorescent Assays on the Barcoded Rapid Assay Platform***

The barcode patterned microchip surface was partitioned into 16 individual microwells using a pre-fabricated PDMS slab. Individual biotinylated PCC candidates were complexed to specific SAC-DNA conjugates, combined into a cocktail, and then self-assembled, via DNA hybridization, onto designated barcode stripes (Figure 2.1). Incubation with a specific concentration of the target protein preceded incubation with a primary capture antibody and then a fluorophore-conjugated secondary detection antibody. During assay execution,



each well represents a different target concentration or assay condition. Once developed, the fluorescence of the barcodes is digitized using a GenePix 4400A array scanner, with an excitation laser power optimized to a power level of 40% (60 W), which maximizes detection sensitivity while also minimizing signal saturation. Data extraction occurs using 10  $\mu\text{m}$  radius circles, taken along the length of a barcode stripe. A fluorescence signal representing the average of all the pixels within a given circle is collected. A total of ten circles (data-blocks) are measured along a 180  $\mu\text{m}$  span of the middle portion for each individual barcode lane in a given well (Figure A2B.22, Appendix 2B – this illustrates the intensity across a stripe compared to the intensity from the data-block extraction). After extraction the data is background corrected. The background signal arises from (a) non-specific binding of the primary and secondary antibodies (independent of [KRas]), but can vary across different barcode stripes), and (b) non-specific binding of KRas protein ([KRas] dependent). Background (a) was assessed by measuring the average signal in the null protein well for each stripe. Background (b) was assessed by measuring the average fluorescence for the dummy ligand (Biotin\* probe) that was in each well. The background-subtracted data was then graphed in Graphpad Prism 7 and fitted to a sigmoidal curve (Hill coefficient=1).

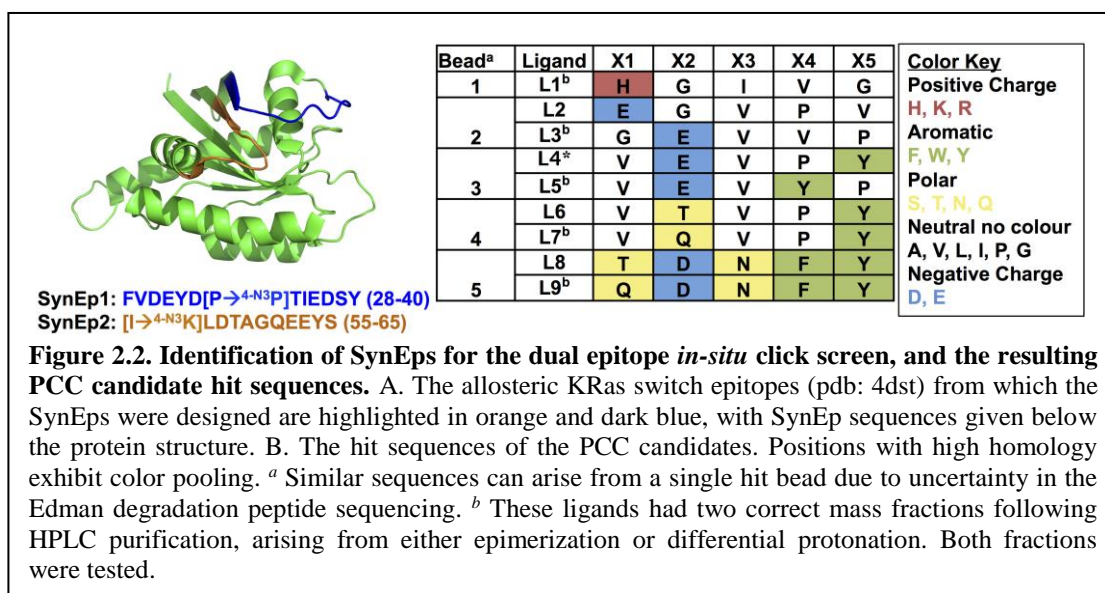
#### ***2.2.6-Measuring the Effect of the Allosteric Ligands on the Intrinsic KRas GTPase Activity***

KRas inhibition assays were carried out using a GTPase Glo Assay kit from Promega (Figure A2B.26). Each candidate inhibitor PCC was initially tested by combining a concentration series of the ligand with 10  $\mu\text{M}$  KRas protein in an opaque white 96-well plate and incubated with 5  $\mu\text{M}$  5'-guanosine triphosphate (GTP) for two hours. The remaining GTP was converted to 5'-adenosine

triphosphate (ATP) over 30 minutes using the reconstituted GTPase Glo reagent before a ten-minute incubation with the detection reagent. Chemiluminescence was measured using a Flexstation 3 plate reader (All wavelengths mode, 500 ms integration), and plotted using Graphpad Prism 7. A full inhibition curve of the most potent inhibitor was then generated using a four-hour incubation with GTP and a 2.5  $\mu$ M to 100  $\mu$ M concentration range. All measurements were done in triplicate.

## Section 2.3-Results and Discussion

### 2.3.1-Optimizing B-RAP Technology Assay Conditions



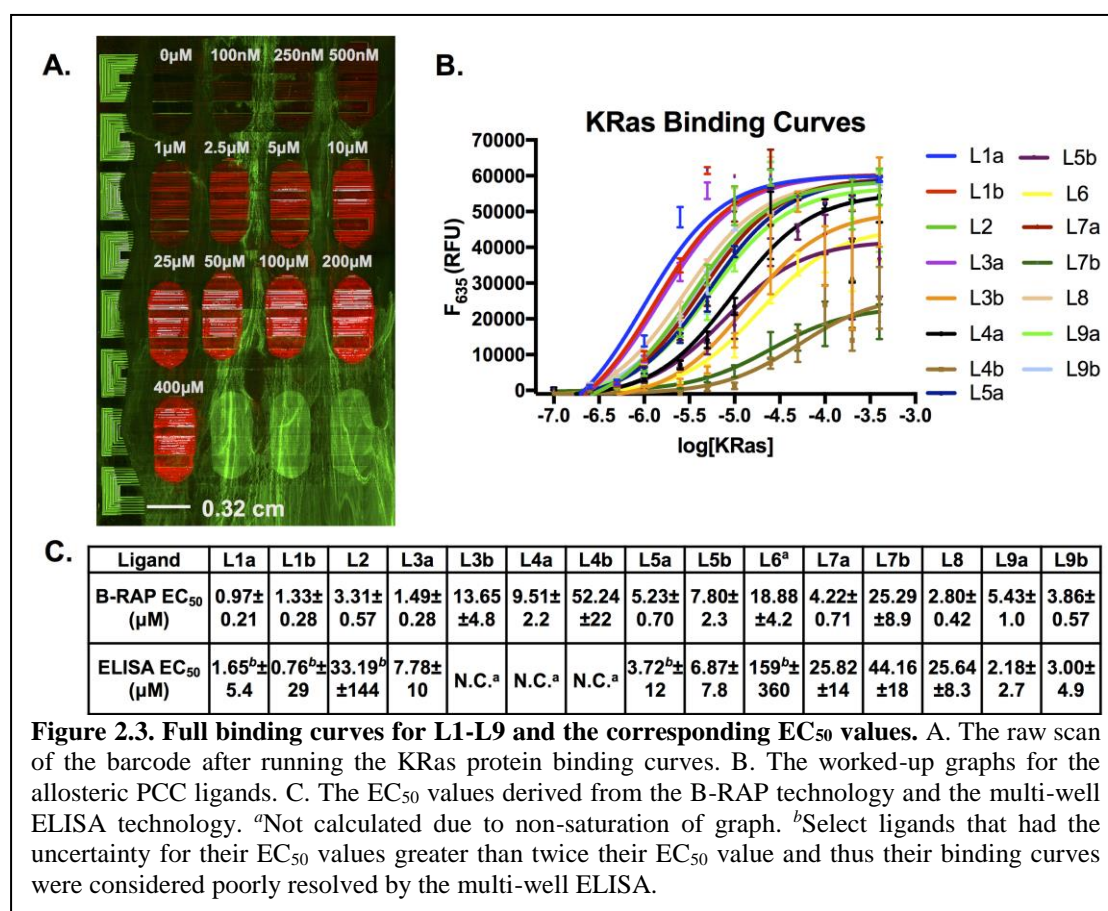
The *in situ* click screen against the Switch I and II KRas protein SynEps (Figure 2.2A) yielded five beads from which nine candidate sequences were determined (Figure 2.2B). Biotinylated candidate ligands were then tested using a single-point IFA with the B-RAP technology (Figure A2B.20, Appendix 2B) to identify appropriate blocking conditions. Modification of the protein incubation solution to include the nonionic surfactant Polysorbate 20 (Tween20) was found to minimize non-specific binding between the KRas protein and the unmodified PLL surface.

### ***2.3.2-Validation of the B-RAP Technology***

Following optimization of the assay conditions the B-RAP technology was subjected to statistical tests to assess the variance in assay results measured within an individual microwell, between microwells on the same chip, and between different microchips. The average percent coefficient of variation (%CV) seen along an individual barcode stripe in the wells above background (500 nM to 400  $\mu$ M KRas) using the values from the data-block extraction method was ~15%. Each microwell contains between two and three full copies of the DNA barcode. For the same barcode lane in the different full barcode sets in the same microwell, the fluorescence output was measured to have an average %CV of ~14% (Figure A2B.21A, Appendix 2B). The %CV between wells on the same microchip run under identical conditions was ~9%. The average %CV for identical barcode lanes between two separate platforms run in parallel by different users was ~18% with an average %CV of ~15% for the 1  $\mu$ M to 400  $\mu$ M range of KRas protein (Figure A2B.21B, Appendix 2B). Additionally, to validate that our data-block extraction method of a portion of the barcode lane captured the  $F_{635}$  for the entire barcode lane, the average  $F_{635}$  from a full-line line scan of the barcode lane was compared to the average  $F_{635}$  resulting from our data-block extraction method. The values from the full-line scan were contained within two standard deviations of the data-block extraction's average  $F_{635}$  (Figure A2B.22, Appendix 2B). This was compared to taking the measurement of individual pixels along the entirety of the barcode lanes in one full set of the 10  $\mu$ M well for one plate (Figure A2B.23A, Appendix 2B) then graphing to find the centroid region (Figure A2B.23B, Appendix 2B), which is the region that is roughly stable in fluorescent

intensity. The average  $F_{635}$ , the standard deviation, and the %CV for each lane was calculated for the full lane, the centroid region of the lane, and the different parts of the centroid region (Table A2C.4, Appendix 2C). The full lane %CVs were in the 20-30% range, while the % CVs of the centroid regions were 10-20%. This arises from edge effects near the microwell walls. Assays of individual PCC candidates (different barcode stripes) collected within a single microwell, and so representing a single point of a binding curve, could be readily distinguished (Table A2C.5, Appendix 2C). These results indicated that the centroid region of a barcode stripe yielded the most reliable data, but also that assay results from different microwells, or different B-RAP chips, could be readily compared.

### 2.3.3-Measuring the $EC_{50}$ of the Allosteric Binding PCC Ligands



After characterizing the B-RAP technology, we used the platform to generate complete binding curves for 15 PCC ligand fractions simultaneously, (Figure 2.3B, **for the binding curves without dummy ligand correction see Figure A2B.24, Appendix 2B**) and determined the EC<sub>50</sub> values for each (Figure 2.3C) (for goodness of fit measurements for the curves see Table A2C.6, Appendix 2C).

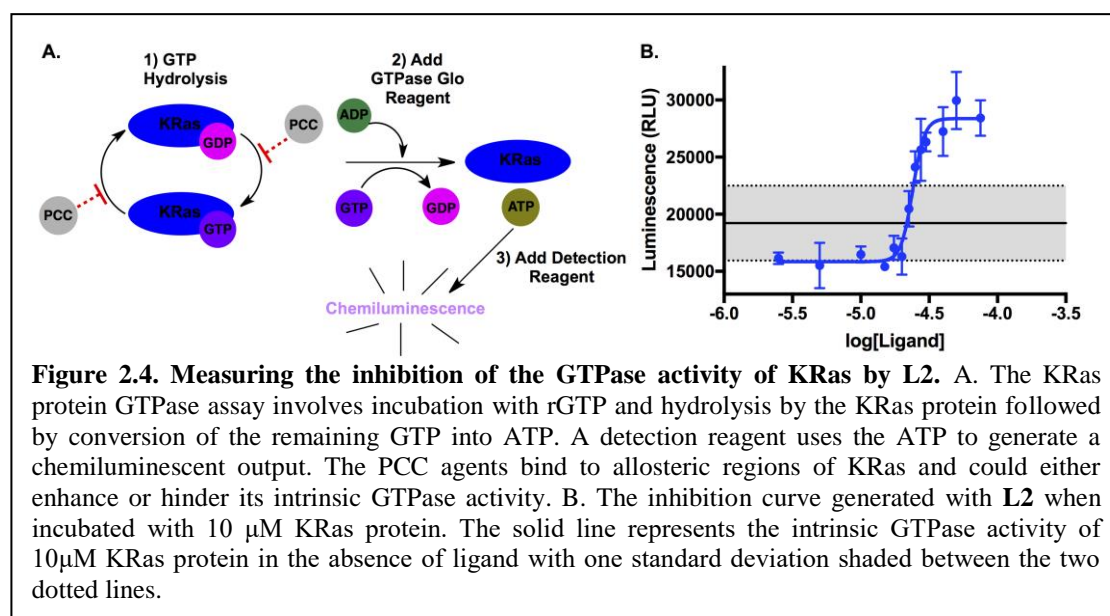
These measurements were comprised of a 13-point concentration series, with each point collected in decaplicate. The EC<sub>50</sub> values enabled the ranking of the ligands, and the best binders were identified to be **L1**, **L2**, and **L8**. The true amino acid sequences for each hit peptide were also distinguished from the artifact sequences that arose from sequencing uncertainties. The true on-bead sequences for the hit beads are identified as **L1**, **L2**, **L5**, **L7**, and **L8**.

We also provide a comparison of the EC<sub>50</sub> values from multi-well ELISA assays (triplicate measurements). While both assays identify **L1a** and **L1b** as the strongest binders, the ELISA assays are significantly noisier, with binding saturation not achieved for several ligands (for the ELISA curves see Figure A2B.25, Appendix 2B). The poor relative performance of the ELISAs arises from a few factors. First, the ligands tested are relatively weak ( $\mu$ M-level) binders, and this exacerbates certain issues associated with the ELISAs. ELISAs are absorbance measurements, and thus have a significantly smaller dynamic range than the B-RAP fluorescence assays. Second, ELISA signal arises from enzymatic amplification, while the B-RAP assays are not amplified. For weak binders, amplified assays tend to be noisy, as both signal and noise are amplified. The improved relative sensitivity and statistics afforded by the B-

RAP technology readily enables the comparative evaluation of these relatively weak KRas binders.

### 2.3.4-Testing the Allosteric Ligands as Inhibitors of KRas GTPase Activity

The ligands identified here were screened for binding to epitopes that exhibit structural fluctuations as the KRas protein switches between its *inactive* 5'-guanosine diphosphate (GDP)-bound form and its *active* GTP-bound form.<sup>25</sup> Consequently, the best three ligand fractions **L1a**, **L2**, and **L8** were probed in a functional, solution phase assay for their ability to disrupt the intrinsic GTPase



enzymatic activity of KRas protein (Figure 2.4A).

This assay measures the enzymatic conversion of GTP to GDP by KRas – a process that can potentially be inhibited. After incubation, an added GTPase Glo™ reagent converts the remaining GTP to ATP, and the ATP is converted into a chemiluminescent signal. Thus, higher chemiluminescence translates to lower KRas enzymatic activity. For the measurements, a fixed [KRas protein] is incubated with varying ligand concentrations. A concentration of 10  $\mu$ M KRas protein was selected after generating a standard curve for the intrinsic

KRas GTPase activity (Figure A2B.26A, Appendix 2B). KRas is a slow acting enzyme, so a KRas/PCC incubation time of two hours was used for the initial survey scans. However, four-hour incubation times were used for the higher resolution data (Figure 2.4B). The survey assays indicated that all three ligands exhibited an inhibitory effect on the KRas protein's GTPase activity, but **L2** was the most potent (Figure A2B.26, Appendix 2B). Thus, the modulation of KRas activity by **L2** was recorded with an expanded concentration range (Figure 2.4B). We found that **L2** switches from weakly *activating* to strongly *inhibiting* above 20  $\mu\text{M}$ . Less than 5% of the rGTP was hydrolyzed in the L2-only (no KRas) wells, and ~61% was hydrolyzed in the KRas-only wells. This result confirms that **L2** lacks any innate GTPase enzymatic activity (Figure A2B.27, Appendix 2B). The sharp transition in the titration curve fits to a Hill coefficient of ~10, and suggests that upon full occupancy of the allosteric switch region, KRas flips into an inactive conformation. An  $\text{IC}_{50}$  value was  $24.0 \pm 1.2 \mu\text{M}$  for **L2**. This is an excellent starting point for a first generation allosteric inhibitor against this challenging target.

## Section 2.4-Conclusions

**Table 2.1. Comparison of the capacity, reagent quantities used, and assay times for multi-well ELISA plates relative to the barcoded rapid assay platform.**

Criteria	ELISA Platform	B-RAP Chip
Full binding curves per Assay	1	15
Relative amount PCC per binding curve (nmol)	7	0.15
Relative amount protein per binding curve (nmol)	300	2.7
Protein concentration points per assay	12	16
Assay run time (h)	10-36	8-10
# Data Points per platform	96	2400

We report on the development and use of a barcoded rapid assay microchip, which allows for the simultaneous evaluation of fifteen PCC candidate ligands in up to sixteen unique assay conditions, with significant associated savings in terms of

both time and reagent use (Table 2.1). In a single day the B-RAP technology was applied to identify the best allosteric KRas binders from a pool of 15 ligands identified from a dual SynEp PCC library *in situ* click screen. The B-RAP technology is designed to yield an equilibrium-based EC<sub>50</sub> value for assessing relative binding strengths. For a number of PCCs, the EC<sub>50</sub> value provides an upper limit for the dissociation constant (K<sub>D</sub>).<sup>1</sup> Importantly, relative binding affinities can provide guidance for selecting ligands for further quantitative characterizations, such as the solution phase KRas activity assay explored here. To this end, the B-RAP technology works well. A comparison of the B-RAP assay metrics relative to standard 96-well plate ELISAs is presented (Table 2.1). Extending this platform to evaluating PCC binders, or other ligand classes, against new protein targets should work well, requiring only an optimization of both concentration ranges (determined by the candidate ligands) and blocking conditions (typically determined by the protein target).<sup>31</sup>

Using the B-RAP platform coupled with the epitope-targeted *in situ* click screening approach, we identified a PCC ligand lead (**L2**) that serves as an allosteric inhibitor of the intrinsic GTPase enzymatic activity of KRas, with an IC<sub>50</sub> value of around 20 μM. **L2** is a first generation ligand, and, as such, can surely be optimized, via medicinal chemistry methods, for increased potency and selectivity. Thus, given the well-known challenging nature of KRas as a drug target, **L2** provides an excellent starting point for developing a more potent inhibitor.

## Section 2.5-References

- (1) Das, S.; Nag, A.; Liang, J.; Bunck, D. N.; Umeda, A.; Farrow, B.; Coppock, M. B.; Sarkes, D. A.; Finch, A. S.; Agnew, H. D.; Pitram, S.; Lai, B.; Yu, M. B.;



- Museth, A. K.; Deyle, K. M.; Lepe, B.; Rodriguez-Rivera, F. P.; McCarthy, A.; Alvarez-Villalonga, B.; Chen, A.; Heath, J.; Stratis-Cullum, D. N.; Heath, J. R. *Angew. Chemie Int. Ed.* **2015**, *54* (45), 13219.
- (2) Farrow, B.; Wong, M.; Malette, J.; Lai, B.; Deyle, K. M.; Das, S.; Nag, A.; Agnew, H. D.; Heath, J. R. *Angew. Chemie Int. Ed.* **2015**, *54* (24), 7114.
- (3) Agnew, H. D.; Rohde, R. D.; Millward, S. W.; Nag, A.; Yeo, W. S.; Hein, J. E.; Pitram, S. M.; Abdul Ahad Tariq, V.; Burns, A. M.; Krom, R. J.; Fokin, V. V.; Barry Sharpless, K.; Heath, J. R. *Angew. Chemie Int. Ed.* **2009**, *48* (27), 4944.
- (4) Yüce, M.; Ullah, N.; Budak, H. *Analyst* **2015**, *140* (16), 5379.
- (5) Jost, C.; Plückthun, A. *Current Opinion in Structural Biology* **2014**, 102.
- (6) Mascini, M.; Palchetti, I.; Tombelli, S. *Angew. Chemie - Int. Ed.* **2012**, *51* (6), 1316.
- (7) Csordas, A. T.; Jørgensen, A.; Wang, J.; Gruber, E.; Gong, Q.; Bagley, E. R.; Nakamoto, M. A.; Eisenstein, M.; Soh, H. T. *Anal. Chem.* **2016**, *88* (22), 10842.
- (8) Kwong, G. A.; Radu, C. G.; Hwang, K.; Shu, C. J.; Chao, M.; Koya, R. C.; Comin-Anduix, B.; Hadrup, S. R.; Bailey, R. C.; Witte, O. N.; Schumacher, T. N.; Ribas, A.; Heath, J. R. *J. Am. Chem. Soc.* **2009**, *131* (28), 9695.
- (9) Bailey, R. C.; Kwong, G. A.; Radu, C. G.; Witte, O. N.; Heath, J. R. *J. Am. Chem. Soc.* **2007**, *129* (7), 1959.
- (10) Boozer, C.; Ladd, J.; Chen, S.; Yu, Q.; Homola, J.; Jiang, S. *Anal. Chem.* **2004**, *76* (23), 6967.
- (11) Kozlov, I. A.; Melnyk, P. C.; Stromborg, K. E.; Chee, M. S.; Barker, D. L.; Zhao, C. *Biopolymers* **2004**, *73* (5), 621.
- (12) Adler, M.; Wacker, R.; Bootlank, E.; Manz, B.; Niemeyer, C. M. *Nat. Methods*

- 2005**, 2 (2), 147.
- (13) Sano, T.; Cantor, C. R. *Proc. Natl. Acad. Sci.* **1990**, 87 (1), 142.
  - (14) Reznik, G. O.; Vajda, S.; Cantor, C. R.; Sano, T. *Bioconjug. Chem.* **2001**, 12 (6), 1000.
  - (15) Ramachandiran, V.; Grigoriev, V.; Lan, L.; Ravkov, E.; Mertens, S. A.; Altman, J. D. *J. Immunol. Methods* **2007**, 319 (1–2), 13.
  - (16) Shin, Y. S.; Ahmad, H.; Shi, Q.; Kim, H.; Pascal, T. A.; Fan, R.; Goddard, W. A.; Heath, J. R. *ChemPhysChem* **2010**, 11 (14), 3063.
  - (17) Cooper, G. M. *Science* **1982**, 217 (4562), 801.
  - (18) Cox, A. D.; Fesik, S. W.; Kimmelman, A. C.; Luo, J.; Der, C. J. *Nature Reviews Drug Discovery*. 2014, pp 828–851.
  - (19) Whitehead, R. P.; Mccoy, S.; Macdonald, J. S.; Rivkin, S. E.; Neubauer, M. A.; Dakhil, S. R.; Lenz, H.-J.; Tanaka, M. S.; Abbruzzese, J. L. *Invest. New Drugs* **2006**, 24, 335.
  - (20) Macdonald, J. S.; Mccoy, S.; Whitehead, R. P.; Iqbal, S.; Wade Iii, J. L.; Giguere, J. K.; Abbruzzese, J. L. *Invest. New Drugs* **2005**, 23, 485.
  - (21) Winquist, E.; Moore, M. J.; Chi, K. N.; Ernst, D. S.; Hirte, H.; North, S.; Powers, J.; Walsh, W.; Boucher, T.; Patton, R.; Seymour, L. *Urol. Oncol.* **2003**, 23 (3), 143.
  - (22) Sharma, S.; Kemeny, N.; Kelsen, D. P.; Ilson, D.; O 'reilly, E.; Zaknoen, S.; Baum, C.; Statkevich, P.; Hollywood, E.; Zhu, Y.; Saltz, L. B. *Ann. Oncol.* **2002**, 13, 1067.
  - (23) Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M. *Nature* **2013**, 503.
  - (24) Sakamoto, K.; Kamada, Y.; Sameshima, T.; Yaguchi, M.; Niida, A.; Sasaki, S.;

- Miwa, M.; Ohkubo, S.; Sakamoto, J. ichi; Kamaura, M.; Cho, N.; Tani, A.  
*Biochem. Biophys. Res. Commun.* **2017**, *484* (3), 605.
- (25) Hall, B. E.; Yang, S. S.; Boriack-Sjodin, P. A.; Kuriyan, J.; Bar-Sagi, D. *J. Biol. Chem.* **2001**, *276* (29), 27629.
- (26) *For detailed protocols of all experiments see Appendix 2A.*
- (27) Yu, J.; Zhou, J.; Sutherland, A.; Wei, W.; Shin, Y. S.; Xue, M.; Heath, J. R.  
*Annu. Rev. Anal. Chem.* **2014**, *7* (1), 275.
- (28) Wei, W.; Shin, Y. S.; Xue, M.; Matsutani, T.; Masui, K.; Yang, H.; Ikegami, S.; Gu, Y.; Herrmann, K.; Johnson, D.; Ding, X.; Hwang, K.; Kim, J.; Zhou, J.; Su, Y.; Li, X.; Bonetti, B.; Chopra, R.; James, C. D.; Cavenee, W. K.; Cloughesy, T. F.; Mischel, P. S.; Heath, J. R.; Gini, B. *Cancer Cell* **2016**, *29* (4), 563.
- (29) Xue, M.; Wei, W.; Su, Y.; Kim, J.; Shin, Y. S.; Mai, W. X.; Nathanson, D. A.; Heath, J. R. *J. Am. Chem. Soc.* **2015**, *137* (12), 4066.
- (30) Boriack-Sjodin, P. A.; Margarit, S. M.; Bar-Sagi, D.; Kuriyan, J. *Nature* **1998**, *394* (6691), 337.
- (31) Gibbs, J. *ELISA Tech. Bull. Corning Inc. Life Sci. Kennebunk, ME* **2001**, No. 3, 1.

## Appendix 2A

### Materials and Experimental Procedures

#### **Table of Contents**

<b>List of Chemicals and Reagents</b>	47
<b>Protocols</b>	
DNA Barcode Chip Patterning and Validation	51
Barcode Validation	52
Expression of Cysteine-Modified Streptavidin (SAC) Protein	53
Isolation of SAC Inclusion Bodies from <i>E. coli</i> Cells	53
Refolding and Purification of SAC Protein	54
Preparation of SAC-DNA Conjugates	54
Biotin Binding Test of SAC-DNA Conjugates	56
WT KRas Protein Expression and Purification	57
<i>In-Situ</i> Library Screen Preclear/Anti-Screen	58
<i>In-Situ</i> Library Screen Product/Target Screens	59
Peptide Synthesis Protocols	60
WT KRas Binding Curves Using the B-RAP Technology	60
WT KRas Binding Curves Using the Multi-Well ELISA Technology	62
Testing Allosteric Ligands for Inhibition of Intrinsic KRas Protein	
GTPase Activity	63
<b>References</b>	64

## Materials

Unless otherwise stated all chemicals were used as received.

### **Barcode Microfabrication and Validation**

Chrome masks of the custom barcode design were purchased from University of California, Los Angeles, Nanoelectronics Research Facility, and a Karl Süss MA/BA6 mask aligner (SÜSS MicroTec AG) was used for UV exposure. Silicon wafers (Wafernet Inc.), SU8-2025, and SU8 developer (Microchem Corp) were used for the barcode mold fabrication. Anhydrous dimethylsulfoxide (DMSO), sodium dodecyl sulfate (SDS), and bis(sulfosuccinimidyl)suberate (BS3) used in barcode fabrication were purchased from American Type Culture Collection (ATCC), Sigma Aldrich, and ThermoFischer Scientific respectively. The Sylgard 184 elastomer, and poly-L-lysine coated glass slides used in DNA barcode microfabrication were purchased from Dow Corning and ThermoFischer Scientific respectively. The poly-L-lysine (PLL) solution (0.1% (w/w)) used for barcode fabrication was purchased from Sigma Aldrich. All ssDNA used for barcode fabrication and barcode validations were purchased from either Bioneer Inc. or IDT Inc.

### **Protein Expression, Purification, and Refolding**

The Bacto Tryptone (Tryptone) and Bacto yeast (yeast) for the preparation of LB broth media were purchased from Becton, Dickinson, and Company. The ampicillin sodium salt, chloramphenicol, and isopropyl  $\beta$ -D-1-thiogalactopyranoside (1,6-IPTG, dioxane free) used for protein expression from *E. coli* were purchased from Sigma Aldrich. The one-shot B21(D3) *E. coli* cells and PQE80 vector (His<sub>6</sub>-tagged human KRas Isoform 4B (residues 1-169)) used for expression of KRas protein were purchased from Life

Technologies and Qiagen respectively. Lysozyme (L6876), DNase I (10104159001), and RNase A (R6513-10MG) used for lysing cells containing cysteine-modified streptavidin (SAC) were purchased from Sigma-Aldrich. Cells containing KRas protein were lysed using a constant pressure cell disruptor (Constant Systems Ltd., Scotland, UK). Surfactants Triton X-100 and polysorbate 20 (Tween20) were purchased from Sigma Aldrich. The 20x phosphate buffered saline with 0.05% Tween 20 (PBST) and phosphate buffered saline (PBS) used for protein purification and immunofluorescent assays (IFAs) were purchased from Cell Signaling Technology and Corning respectively. The sodium bicarbonate ( $\text{NaHCO}_3$ ), ammonium acetate ( $\text{NH}_4\text{OAc}$ ), sodium acetate ( $\text{NaOAc}$ ), sodium chloride ( $\text{NaCl}$ ), imidazole, tris(hydroxymethyl)aminomethane (Tris), tris(hydroxymethyl)aminomethane hydrochloride salt ( $\text{Tris}\cdot\text{HCl}$ ), guanidinium chloride ( $\text{Guan}\cdot\text{HCl}$ ), magnesium chloride pentahydrate ( $\text{MgCl}_2\cdot 5\text{H}_2\text{O}$ ), and beta-mercaptoethanol ( $\beta\text{ME}$ ) used in protein purification and IFA assays were purchased from Sigma Aldrich. The 2-aminobiotin agarose resin, Superdex 75 (10/300) increase column, and Ni-NTA superflow cartridge used for fast protein liquid chromatography (FPLC) purification were purchased from Sigma Aldrich, GE Healthcare Life Sciences, and Qiagen respectively. The Amicon Ultra-15 and Ultra-4 centrifugal filters used to concentrate protein samples were purchased from EMD Millipore.

### **SAC-DNA Conjugation and Validation**

The tris-(2-carboxyethyl)phosphine hydrochloride (TCEP), anhydrous N,N-dimethylformamide (DMF), N-succinimidyl-4-formyl benzaldehyde (S-4FB) and maleimide 6-hydrazino-nicotinamide (MHPH) used for the conjugation of ssDNA to cysteine-modified streptavidin (SAC) were purchased from Sigma Aldrich and Solulink.

The biotin-A<sub>20</sub>-Cy3 (Biotin\*) probe used to test the biotin binding ability of the SAC-DNA set and used as a biotinylated blank for IFA assays was purchased from IDT Inc. The complementary ssDNA' used for conjugation to SAC were purchased from Bioneer Inc.

### **In Situ Library Screen and Hit Bead Sequencing**

The mouse anti-biotin-alkaline phosphatase conjugated antibody (ab) (#A6561), goat anti-rabbit-alkaline phosphatase conjugated ab (#A8025), and rabbit anti-Ras ab (CST #3965) used for the combined anti screen/pre-clear and the subsequent product/target screens were purchased from Sigma Aldrich and Cell Signaling Technology respectively. The combined 5-bromo-4-chloro-3-indoyl phosphate (BCIP)/ nitro blue tetrazolium (NBT) (#S3771) used to develop hits during the library screens was purchased from Promega. The concentrated hydrochloric acid used to quench the BCIP/NBT development was purchased from Sigma Aldrich. Sequencing of bead hits occurred via Edman degradation sequencing on a Procise Protein Sequencer (Applied Biosystems, California).

### **Peptide Synthesis and Purification**

Fmoc-protected amino acids were purchased from Anaspec, AAPTec, Bachem, ChemPep, and Sigma-Aldrich. Biotin NovaTag™ resin was obtained from EMD Chemicals, Inc. and used for the synthesis of biotinylated peptides and epitopes used for the screens using standard Fmoc/'Bu coupling and cleavage protocols. The peptide one-bead-one compound (OBOC) library was prepared on Tentagel Resin purchased from RAPP Polymere. The Fmoc-protected propionic acid polyethylene glycol (PEG<sub>n</sub>) linkers were purchased from ChemPep Inc. The L-ascorbic acid and copper (I) iodide (CuI) used for

click reactions were purchased from Sigma Aldrich. The N-methyl pyrrolidine (NMP), 1-[Bis (dimethylamino) methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU), and *N,N'*-diisopropylethylamine (DIPEA) used in peptide synthesis were bought from EMD Chemicals, Inc., ChemPep, and Sigma-Aldrich respectively. Piperidine, trifluoroacetic acid (TFA), and triethylsilane (TESH) were purchased from Sigma-Aldrich. The diethyl ether used to precipitate crude peptide was purchased from JT Baker. The Omnisolv grade acetonitrile (MeCN) used for peptide purification was purchased from EMD Millipore. Unless otherwise stated, peptide preparation was performed using a Titan 357 Automatic Peptide Synthesizer (AAPPTec, Louisville, KY) or a Liberty 1 Automated Peptide Synthesizer (CEM, North Carolina). Mass analysis was performed using a Voyager De-Pro matrix assisted laser desorption ionization time-of-flight mass spectrometer (MALDI-TOF MS) (Applied Biosystems, California). The crude peptides were dissolved in either DMSO (Sigma Aldrich) or (1:1) MeCN/doubly distilled water (MQ H<sub>2</sub>O) w/ 0.1% TFA before purification by a gradient of 0% to 50% acetonitrile in MQ H<sub>2</sub>O with 0.01% (v/v) TFA using a RP-HPLC (Beckman Coulter System Gold 126 Solvent Module and 168 Detector) using a C18 reversed phase semi-preparative column (Phenomenex Luna 10 µm, 250 × 10 mm). The concentration of peptides and epitopes was determined using a Nanodrop 2000 Spectrophotometer (ThermoFischer Scientific Inc., Massachusetts).

### **B-RAP Immunofluorescent Assays and Multi-Well Enzyme-Linked Immunosorbent Assays**

The Bovine Serum Albumin (BSA, Biotin free A1933-25G) used in the IFAs and multi-well enzyme-linked Immunosorbent assays (ELISAs) was purchased from Sigma-Aldrich.



The non-fat dry milk powder used in the enzyme-linked Immunosorbent assays (ELISAs) was purchased from Best Value. The rabbit anti-Ras (CST #3965), Goat anti-rabbit IgG HRP-linked (CST #7074), goat anti-rabbit HRP-linked (CST #7074), and goat anti-rabbit-Alexafluor 647 conjugated (ab150079) were purchased from Cell Signaling Technologies and Abcam respectively. The ELISAs were run on either 96-well clear Pierce Neutravidin Plates (#15129) or Pierce Neutravidin Coated Plates (#15127) purchased from ThermoFischer Scientific. The TMB Microwell Peroxidase Substrate System (#50-76-00) that was used to develop ELISAs was purchased from KPL. The sulfuric acid ( $\text{H}_2\text{SO}_{4(\text{aq})}$ ) used to quench the enzymatic amplification reaction in the ELISAs was purchased from JT Baker. The 96-well ELISA plates were read using a Flexstation 3 plate reader (Molecular Devices LLC, Sunnyvale, CA). All barcode slides were scanned using an Axon GenePix 4400A (Molecular Devices LLC, Sunnyvale, CA).

### **Measuring the Functional Effect of the Allosteric Ligands on KRas Protein GTPase Activity**

The intrinsic GTPase activity of WT KRas protein was measured using the GTPase-Glo Assay Kit (#V7681) from Promega Corporation (Madison, WI) on opaque white 96-well plates (#6005290) from Perkin Elmer Life Sciences (Waltham, MA). Luminescence was recorded on the Flexstation 3 plate reader used for multi-well ELISAs.

## **Procedures**

### ***DNA Barcode Chip Patterning and Validation***

The DNA barcode chips were prepared by micro channel-guided flow patterning as described in **References 27 and 28** of the main text. A PDMS slab having the micro-channels was made by soft lithography on a silicon wafer. Its mold was designed as

Figure S1 and prepared with SU8 2025 negative photoresist. The fabricated mold contained microfluidic circuits of 20 parallel channels with 50  $\mu\text{m}$  width and  $\sim 40$   $\mu\text{m}$  height. Sylgard® 184 PDMS pre-polymer and curing agents were mixed in a 10:1 ratio, degassed,  $\sim 60\text{g}$  of the mixture poured onto the mold, and baked for two hours at 80 °C for curing. The cured PDMS slab was peeled off the mold, cut into individual microfluidic molds, and the inlet and outlet holes of the microfluidic circuits were punched with the sizes of two mm and 0.5 mm respectively. The number of the inlets and outlets punched out were determined by the number of single stranded DNAs (ssDNAs) used in the assay, and fifteen orthogonal ssDNAs (**B-Q**, Table S2) were used in this study. The slab was then aligned with a PLL glass slide, and bonding occurred with baking at 80 °C for two hours. After cooling briefly, the inlet wells were loaded with 3  $\mu\text{L}$  of a PLL solution (0.1% ( $\text{m}/\text{m}$ ) in  $\text{H}_2\text{O}$ ), and the PLL solution was flowed and dried by 13.8 kPa nitrogen gas blowing through the solution-loading device overnight. The next day, C6 amine-modified DNA solutions (300  $\mu\text{M}$  in (3:2 ( $\text{v}/\text{v}$ )) PBS/DMSO) were individually mixed (1:1) with a 2 mM BS3 cross-linker solution in PBS. Each freshly prepared mixture was flown through a channel under 13.8-20.6 kPa of nitrogen gas using the solution-loading device for 1 hour, and then only the assembled PDMS slab and the bonded PLL slide were incubated at room temperature for 2 hours in a humidified chamber. After incubation, the PDMS slab was removed, and the DNA patterned PLL slides were washed both a 0.02% aqueous SDS solution and doubly distilled water (MQ  $\text{H}_2\text{O}$ ) (3x), and spun dry.

### ***Barcode Validation***

To validate the DNA barcode chips, a 5'-modified Cy3-labeled complementary ssDNA cocktail was prepared in 1% BSA in PBS (50nM each ssDNA). The validation occurred over two rounds (B, D, F, H, K, N, P, M then C, E, G, I, L, O, Q) in order to check for channel leaks and crossover. A 120  $\mu$ L aliquot of the validation solution was applied to a small region at the bottom edge of the DNA barcode before incubating at 37 °C for one hour. After incubation, this region was washed with 1% BSA in PBS followed by PBS (2x), and the slide was spun dry before being scanned by Axon GenePix 4400A (532 nm, PMT 450, Power 15% (23W)) (Figure S2).

### ***Expression of Cysteine-Modified Streptavidin (SAC) Protein***

The SAC protein was expressed using a modification of the procedure reported by Sano and Cantor.<sup>1</sup> A 100 mL starter culture of autoclaved LB media (10.0 g Tryptone, 5.00 g yeast, 10.0 g NaCl per L H<sub>2</sub>O) was prepared by inoculating with 50  $\mu$ L of 100 mg/mL of ampicillin (final concentration 50  $\mu$ g/mL) and 100 $\mu$ L of 34 mg/mL chloramphenicol (final concentration 34  $\mu$ g/mL) followed by a sterile pipet scraping of a 50% (v/v) glycerol stock containing transformed *E. coli* BL21(D3) cells. The starter culture incubated overnight at 37 °C and 250 RPM before adding 10.0 mL of starter culture aliquots to six 2800mL Fernbach-Style Culture Flasks containing 1.00 L autoclaved LB media with 500  $\mu$ L of 100 mg/mL of ampicillin (final concentration 50  $\mu$ g/mL), 1000  $\mu$ L of 34 mg/mL chloramphenicol (final concentration 34  $\mu$ g/mL), and 1000 $\mu$ L of 40% (w/w) autoclaved glucose (final concentration 0.4% (w/w)). The flasks were left to culture at 37.0 °C, 250 RPM until A<sub>680</sub> = 0.500, and induction was triggered with 1000  $\mu$ L of a 400 mM 1,6-IPTG solution (final concentration 400  $\mu$ M). Expression continued at 37.0 °C, 250 RPM for four hours before spinning down the cells at 6000 RPM, 5 minutes at 4 °C. The cells

were resuspended in 50 mL of a 10 mM Tris, 1 mM EDTA, 130 mM NaCl buffer at pH=8.0 and spun down (2x). The cells were then flash frozen in N<sub>2(l)</sub> and stored at -80.0 °C until needed.

### ***Isolation of SAC Inclusion Bodies from e. Coli Cells***

The cell pellet was thawed in ice before resuspending in two 50-mL falcon tubes with 40 mL of TEX buffer (30mM Tris, 2mM EDTA, 0.1% TritonX). Each tube was charged with 40 mg fresh lysozyme powder (Final concentration 1.0 mg/mL), vortexed until mixed, and allowed to lyse for 30min while tumbling at RT. The solution was very viscous after lysis. The DNA and RNA were degraded by adding 400 µL of 10 mg/mL DNase and 10 mg/mL RNase in TE Buffer (10 mM Tris, 130 mM NaCl, 1 mM EDTA) (final concentration 10 µg/mL), 960 µL of 500 mM MgCl<sub>2</sub> (final concentration 12 mM), and 40 µL of 1 M MnCl<sub>2</sub> (final concentration 1 mM) to each tube of cell lysate, and the solution was allowed to digest for 30 minutes while tumbling at RT. After digestion, the solution was spun down at 7800 RPM, RT for 10 minutes. The resulting inclusion body (IB) pellets were both washed in 40 mL TEX buffer and spun down at 7800 RPM, 5 minutes at RT. Pellets were washed with 40 mL buffer minus Triton X again before spinning down at 7800 RPM, 5 minutes at RT once more. Each pellet was taken up in 10 mM Tris and spun down at 7800RPM, 10 minutes at RT, aliquotted, and stored at -80.0 °C until needed. If the final pellet is light brown then some DNA is still present. This will be removed at the beginning of the refolding procedure.

### ***Refolding and Purification of SAC Protein***

The procedure described here is a modification of the procedure developed by Sano and Cantor.<sup>1</sup>

**\*\*\*After the initial denaturing keep all solutions at 4 °C\*\*\***

An IB aliquot was dissolved in 1000  $\mu$ L denaturing buffer (6 M Guanidine • HCl at pH=1.5 with 10 mM  $\beta$ ME), vortexed, spun down at 13,000 RPM, 2 min at RT, and filtered using a 0.45  $\mu$ M low-protein binding filter. **The resulting solution should be clear and nearly colorless.** The  $A_{280}$  was measured on a Nanodrop2000 spectrophotometer, and the concentration of denatured SAC monomer was calculated.<sup>2</sup> The denatured SAC solution was diluted to 1000  $\mu$ L in denaturing buffer and added dropwise to a rapidly stirring solution of refolding buffer (50 mM  $\text{NH}_4\text{OAc}$ , 150 mM NaCl, and 10 mM  $\beta$ ME at pH=6.0) (Final [denatured SAC]  $\sim$ 4  $\mu$ M). The stir rate was then decreased to about half of its original value, and the solution was covered by aluminum foil to refold overnight. After sterile filtration with a 0.45  $\mu$ m low-protein binding filter the resulting solution was concentrated to 10-15 mL using Amicon Millipore filters (10,000-30,000 MWCO) before dialyzing the refolded SAC protein in buffer A (50 mM  $\text{NaHCO}_3$ , 500 mM NaCl, 10 mM  $\beta$ ME at pH=11.0) until the solution had a pH of  $\sim$ 11 (about 2 hours). The crude protein was then diluted (1:1) with buffer A, mixed with 2 mL of 2-iminobiotin agarose resin, and allowed to incubate with tumbling in the cold room for one hour. After incubation the supernatant was eluted (3x) before eluting with buffer A until the  $A_{280}$  went to baseline. Pure SAC was eluted with buffer B (50 mM  $\text{NaOAc}$ , 50 mM NaCl at pH=4.0) until the  $A_{280}$  went to baseline again. Fractions with pure SAC were pooled and dialyzed against a PBS solution (PBS, 10 mM  $\beta$ ME, pH=7.5) overnight. The SAC was concentrated to  $\sim$ 1  $\text{mg/mL}$  final concentration, divided into 100  $\mu$ L aliquots, and stored at -80 °C.

### ***Preparation of SAC-DNA Conjugates***

For each planned SAC-DNA conjugation, two Zeba columns were prepared (3 x 300  $\mu$ L of 5 mM TCEP in PBS, 3.9k RPM, 1 min at RT). Each 100  $\mu$ L aliquot of SAC was desalted in two separate Zeba columns to remove the  $\beta$ ME (3.9k RPM, 2 min at RT). After transferring to Eppendorf tubes, 6  $\mu$ L anhydrous DMF was added followed by 6  $\mu$ L MHPH (100 mM in anhydrous DMF). Separate Eppendorf tubes were charged with 80  $\mu$ L of 500  $\mu$ M of conjugation ssDNA in PBS followed by 15  $\mu$ L anhydrous DMF and 20  $\mu$ L S-4FB (100 mM in anhydrous DMF). The SAC and DNA solutions were vortexed gently, briefly spun down, and left to react at RT in the dark for four hours. For each conjugation in progress, four Zeba columns were buffer exchanged with citrate buffer (150 mM NaCl, 50 mM sodium citrate, pH=6.0) (3 x 300  $\mu$ L citrate buffer, 3.9k RPM, 1 min at RT). The SAC and DNA solutions were desalted separately (2 x 3.9k RPM, 2 min, at RT) before combining each SAC protein aliquot with a unique ssDNA solution. The solutions were vortexed gently, briefly spun down, and left to react in the dark at RT overnight. The reactions were quenched by placing at 4 °C. Each SAC-DNA conjugate was purified by FPLC using a Superdex75 Increase column (isocratic in PBS, 0.5 mL/min, 0.5 mL fractions, 75 minutes). Fractions containing pure SAC-DNA were pooled and concentrated using Amicon Ultra-4 Centrifugal filters (30k MWCO): 3900 RPM, 30 minutes at 4 °C. The concentrated SAC-DNA proteins were quantified<sup>3</sup> using a Nanodrop2000 spectrophotometer in the ssDNA nucleic acid mode (using two for the average number of ssDNA strands conjugated as previously established) and stored at 4 °C until needed.

### ***Biotin Binding Test of SAC-DNA Conjugates***

#### **Buffers used:**

Wash buffer: PBS + 0.05% Tween20 (PBST)

Blocking Buffer: PBS + 1% BSA

**\*Wash steps used 50  $\mu\text{L}$ /well**

**\*Incubation steps used 30  $\mu\text{L}$ /well**

**\*\*After loading the Biotin\* probe, change pipette tips every time that you aspirate or add solution to a well to prevent cross contamination\*\***

A prefabricated PDMS template was aligned onto the DNA barcode and the microchip slide was taped into a 10 cm petri dish. The wells were washed with PBST before loading blocking buffer and placing the platform into a 37 °C incubator for one hour. A cocktail containing 50 nM of each SAC-DNA in PBS was prepared and added to the pre-blocked wells. The SAC-DNA conjugates were allowed to hybridize to the DNA barcode at 37 °C for one hour before washing the wells with PBST (3x). Each well was loaded with 50 nM, 100 nM, 150 nM, 200 nM, 300 nM, or 400 nM Biotin\* in PBS (Figure S5), and the platform was left to shake covered at RT for one hour. The wells were washed with PBST (3x) before peeling off the PDMS slab and dipping the barcode into PBS, (1:1) PBS/MQ H<sub>2</sub>O, MQ H<sub>2</sub>O (2x). The barcode was then spun dry and read on the Genepix (532 nm, PMT 450, Power 15% (23 W)).

#### ***WT KRas Protein Expression and Purification***

The KRas protein was expressed and purified using a modification of the procedure reported by Kuriyan.<sup>4</sup> A starter culture of 100 mL of autoclaved LB media was inoculated with 100  $\mu\text{L}$  of 100  $\text{mg}/\text{mL}$  of ampicillin (final concentration 100  $\mu\text{g}/\text{mL}$ ) followed by a scraping of a 25% (v/v) glycerol stock containing transformed *E. coli* BL21(D3) cells. The starter culture was left in an incubator at 37.0 °C, 250 RPM overnight before adding 10.0

mL starter culture aliquots to six 2800 mL Fernbach-Style Culture Flasks containing 1.00 L autoclaved LB media with 1000  $\mu$ L of 100  $\text{mg}/\text{mL}$  of ampicillin (final concentration 100  $\text{mg}/\text{mL}$ ). The flasks were left to culture at 37.0  $^{\circ}\text{C}$ , 250 RPM until  $A_{680} = 0.500\text{-}0.600$  and induction was triggered with 1000  $\mu$ L of a 250 mM 1,6-IPTG solution (final concentration 250  $\mu$ M). The cells were then left to express overnight at 18.0  $^{\circ}\text{C}$ , 250 RPM before being spun down, resuspended in buffer A (20 mM Tris, 500 mM NaCl, 20 mM imidazole, 5 mM  $\text{MgCl}_2$ , pH=8.0), flash frozen in  $\text{N}_{2(l)}$ , and stored at -80.0  $^{\circ}\text{C}$  until needed. After thawing and douncing, the cells were lysed using a cell disruptor, the cell wall lysate spun down at 8000 RPM, 4  $^{\circ}\text{C}$  for 20 minutes, sterile filtered with a 0.45  $\mu$ m low-protein binding filter, and purified using FPLC with a Ni-NTA superflow cartridge and a gradient of buffer A to buffer B (20 mM Tris, 300 mM NaCl, 250 mM imidazole, 5 mM  $\text{MgCl}_2$ , pH=8.0). Fractions containing pure KRas were pooled and dialyzed against Tris buffered saline (TBS) (25 mM Tris, 150 mM NaCl, 10 mM  $\text{MgCl}_2$ , pH=7.5) overnight. The resulting solution was concentrated using Amicon Ultra-15 centrifugal filters (10k MWCO), quantified,<sup>2</sup> separated into aliquots, flash frozen in  $\text{N}_{2(l)}$ , and stored at -80.0  $^{\circ}\text{C}$  until needed.

### ***In-Situ Library Click Screen Combined Preclear/Anti-Screen***

The *in-situ* click dual SynEp library screen followed a procedure similar to the one outlined in **Reference 1** from the main text using 450 mg of Pra-capped one-bead-one-compound (OBOC) library. Blocking was performed overnight at 4  $^{\circ}\text{C}$  with blocking buffer (1% BSA and 0.1% Tween20 in TBS). After washing with blocking buffer (3 x 3 minutes) incubation with 25  $\mu$ M of each scrambled SynEp in binding buffer (0.1% BSA and 0.1% Tween20 in TBS) occurred overnight at 4  $^{\circ}\text{C}$ . The library was washed with



TBS (3 x 1 minute) then stripped with 7.5 M Gua•HCl (pH=2.0) at RT for one hour to remove any non-covalently bound scrambled SynEps. Ten rinses with TBS preceded another incubation with blocking buffer at RT for one hour. After five quick rinses of the library with blocking buffer the library was incubated with a cocktail of a (10,000:1) dilution of the mouse anti-biotin-alkaline phosphatase conjugated ab (#A6561), (1,000:1) dilution of the rabbit anti-Ras ab (CST #3965), and a (10,000:1) dilution of the goat anti-rabbit-alkaline phosphatase ab (#A8025) in binding buffer to perform the preclear and anti-screen in one assay. Wash the beads (5 x 3 minutes) with a high salt buffer (25 mM Tris•HCl, 10 mM MgCl<sub>2</sub>, 700 mM NaCl, pH=7.4) and a low salt buffer (5 x 3 minutes) (25 mM Tris•HCl, pH=7.4). The developing buffer was prepared with 66 µL of BCIP (50 mg/mL in 70% DMF) in 10mL of developing buffer (100 mM Tris•HCl, 150 mM NaCl, 1 mM MgCl<sub>2</sub>) and incubated with the library beads in a 20cm petri dish for ten minutes before adding 66 µL of NBT (50 mg/mL in 70% DMF) and incubating for an additional fourteen minutes. The beads were then washed 5x with TBS, and stored in 0.1 M HCl<sub>(aq)</sub> in a 20 cm petri dish. Any beads that turned purple during the combined preclear/anti-clear were promiscuous binders and consequently were picked out using a 10-µL micropipette and discarded. After removing all of the sticky beads the remaining beads were washed with 7.5 M Guan•HCl (pH=2.0) for 30 minutes, rinsed with MQ H<sub>2</sub>O (10x), and incubated in NMP overnight to remove any trace purple coloring. Final rinses with MQ H<sub>2</sub>O (3x), TBS (7x) preceded an overnight incubation at 4 °C with blocking buffer.

#### ***In-Situ Library Click Screen Product/Target Screens***

The pre-blocked library was washed with blocking buffer (3 x 5 minutes) before loading 25 µM of each SynEp in binding buffer and incubating at RT overnight. After

rinsing with TBS (3x) the library was incubated with 7.5 M Guan•HCl (pH=2.0) for one hour at RT before rinsing with TBS (10x). The library then underwent an additional one hour incubation with blocking buffer at RT before rinsing with blocking buffer (5x), and incubating with a (10,000:1) dilution of the mouse anti-biotin-alkaline phosphatase conjugated ab in binding buffer for one hour at RT. Development of the library followed the same procedure as the preclear/anti-screen, and the darkest beads were set aside for Edman degradation sequencing. The remaining ~50 light purple beads from the product screen were prepped following the same procedure after the preclear/anti-screen and screened again, using appropriately scaled amounts of reagents, against 25  $\mu$ M of the full-length KRas protein. After developing, additional beads were picked for a total of seven dark beads from the product/target screens of which five beads yielded readable sequences by Edman degradation sequencing.

### ***Peptide Synthesis Protocols***

All cyclic peptides and epitopes were prepared following the procedures outlined in **Reference 1**.

The peptides and epitopes were isolated using the following procedure. The resin was rinsed with DCM (5x) and dried under vacuum. A 20 mL scintillation vial was charged with a stir-bar, resin, and 3-5 mL cleavage solution (95% TFA, 2.5% TESH, 2.5% H<sub>2</sub>O) and allowed to stir at room temperature for 2-2.5 hours. The solution was then filtered into 40 mL of cold diethyl ether, vortexed for 10 seconds, and stored at 4 °C overnight. The precipitated protein was centrifuged into a pellet at 4500 RPM for 10-15 minutes prior to decantation of the supernatant. The crude peptides were dissolved in either DMSO or (1:1) MeCN/H<sub>2</sub>O w/ 0.1% TFA before HPLC purification, and lyophilization

of desired fractions. The resulting lyophilized powder was dissolved in DMSO, quantified,<sup>2</sup> and stored at 4 °C when not in use.

***WT KRas Binding Curves Using the B-RAP Technology***

***Buffers used:***

Wash buffer: PBS + 0.05% Tween20 (PBST)

Blocking Buffer: PBS + 1% BSA

Protein Incubation Buffer: Tris-buffered saline (TBS) + 0.05% Tween20 (TBST)

1° ab buffer: PBS + 5% BSA

2° ab buffer: PBS + 1% BSA

**\*Wash steps used 50  $\mu\text{L}$ /well**

**\*Incubation steps used 30  $\mu\text{L}$ /well**

**\*\*The plate must be covered during incubation steps to protect the fluorescent blank\*\***

**\*\*After loading the KRas protein change tips every time that solution is aspirated or added to a well to prevent cross-contamination\*\***

A pre-fabricated PDMS template was aligned onto the DNA barcode microchip, and the microchip was taped into a 10 cm petri dish. The wells on the platform were wet with 50  $\mu\text{L}$  PBST before filling with blocking buffer and placing into a 37 °C incubator for 1hr. Concurrently, 40  $\mu\text{L}$  1% BSA in PBS solutions containing 750 nM of a SAC-DNA conjugate and 3.75  $\mu\text{M}$  of one biotinylated PCC ligand or biotin-A<sub>20</sub>-Cy3 blank were prepared for each SAC-DNA conjugate. The biotinylated ligands were allowed to complex with the SAC protein for one hour before pooling the SAC-DNA-ligand solutions (final [SAC-DNA-ligand conjugates] = 50 nM). The blocking buffer was

aspirated, and each well was loaded with the SAC-DNA-ligand conjugates cocktail for hybridization with the DNA barcode at 37 °C for one hour. The wells were washed with PBST (3x) before loading serially diluted solutions of KRas protein in protein buffer (0 to 400 µM). After shaking at RT for one hour, the wells were rinsed with PBST (5x), making sure to pipet up/down with the first addition of PBST. A (100:1) dilution of CST rabbit anti-Ras Ab (#39655) in 1° ab buffer was added to each well before shaking at RT for one hour. After rinsing the wells with PBST (3x), the wells were loaded with a (200:1) dilution of Abcam goat anti-rabbit-Alexafluor 647 linked ab (ab150079) in 2° ab buffer before shaking at RT for one hour. A final rinse of the wells with PBST (3x) preceded peeling off the PDMS slab from the barcode microchip and dipping the barcode into the following solutions: PBS, (1:1) PBS: MQ H<sub>2</sub>O, MQ H<sub>2</sub>O (2x). After being spun dry, the barcode was read on the Genepix (635 nm, PMT 600, PWR 40% (60 W); 532 nm, PMT 450, PWR 15% (23 W)). Data was extracted using  $10^{\text{data blocks}/\text{barcode lane}}$ , double background corrected using the ab-only well fluorescence and dummy ligand fluorescence in each well, and graphed in Graphpad (Sigmoidal 4PL mode with the Hill coefficient set=1). The peeled-off PDMS slab was rinsed under MQ H<sub>2</sub>O and stored in MQ H<sub>2</sub>O at RT until further use.

### ***WT KRas Binding Curves using the Multi-Well ELISA Technology***

#### **Buffers used:**

- Blocking Buffer: TBS + 5% milk + 0.05% Tween20
- Antibody (ab) Buffer: TBS, 5% BSA, 0.05% Tween20
- Binding Buffer: TBS, 0.1% BSA, 0.05% Tween20

**\*All steps were completed at room temperature**

**\*All wash steps used 200  $\mu\text{L}$  solution/well**

**\*All incubations used 100  $\mu\text{L}$  solution/well except for the 5% milk blocking step, which used 200  $\mu\text{L}$  solution/well**

A 96-well Pierce Neutravidin Plate was washed with binding buffer (3 x 5 minutes at RT) before loading plate with a 1  $\mu\text{M}$  solution of either blank (biotin-PEG<sub>5</sub>-NHAc) (singly) or biotinylated PCC ligand (in triplicate). The plate incubated for two hours before washing with binding buffer (3 x 5 minutes). Blocking buffer was added to each well and the plate blocked for one hour before undergoing washing with binding buffer (3 x 5 minutes). Each well was loaded with either binding buffer or KRas solution (0  $\rightarrow$  300  $\mu\text{M}$ ), and the plate was incubated for thirty minutes. Plate washing with binding buffer (3 x 5 minutes) preceded incubating the plate with a (1000:1) dilution of 1<sup>o</sup> antibody (ab) rabbit anti-Ras (CST #3965) in ab buffer for thirty minutes. The plate was washed with binding buffer (3 x 5 minutes), loaded with a (2000:1) dilution of 2<sup>o</sup> ab goat anti-rabbit, HRP-linked ab (CST #7074) in ab buffer, and incubated for an additional thirty minutes. The plate was rinsed with binding buffer (3 x 5 minutes), TBS (1 x 5 minutes), loaded with a (1:1) mixture of TMB Peroxidase Solution and TMB Peroxidase Solution B, and developed with occasional agitation for 8-12 minutes. After quenching the enzymatic reaction with 1M H<sub>2</sub>SO<sub>4(aq)</sub> (100  $\mu\text{L}$ ) the plate was read at  $\lambda = 450$  nm within ten minutes. The data was double background corrected using the ab-only absorbance and the dummy ligand absorbance, plotted using Prism GraphPad 7 (Sigmoidal 4PL mode with the Hill coefficient set=1), and an EC<sub>50</sub> value was calculated.

#### ***Testing Allosteric Ligands for Inhibition of Intrinsic KRas Protein GTPase Activity***

The GTPase assays were run in triplicate on a multi-well plate using the GTPase Glo Assay kit from Promega with 10  $\mu\text{M}$  KRas protein and varying concentrations of ligand (1  $\mu\text{M}$  to 100  $\mu\text{M}$  **L1a**, **L8** and 2.5  $\mu\text{M}$  to 250  $\mu\text{M}$  **L2**). All reagents were warmed to RT before use. A single opaque white 96-well plate was charged with 12.5  $\mu\text{L}$  GTPase/GAP buffer (GTPase buffer w/ 1 mM DTT), 10  $\mu\text{M}$  KRas in GTPase/GAP buffer, or 10  $\mu\text{M}$  KRas protein with either **L1a**, **L2**, or **L8**. **Running the survey assays on the same plate**

**allowed for direct comparison of the curves, but it necessitated the use of the first row/column on the multi-well plate which introduced some noise to the low ligand concentration points.** A 12.5  $\mu\text{L}$  aliquot of a 2x GTP solution (10  $\mu\text{M}$  rGTP in GTPase Buffer) was then added to each well before allowing the plate to shake at RT for two hours (initial GTPase inhibition assay). The GTPase Glo reagent was reconstituted in the GTPase Glo Buffer (4  $\mu\text{L}$  GTPase Glo (500x) reagent, 1996  $\mu\text{L}$  GTPase Glo Buffer, 1  $\mu\text{L}$  10 mM ADP) immediately before adding 25  $\mu\text{L}$  to each well. Shaking at RT for thirty minutes preceded the addition of 50  $\mu\text{L}$  of the detection reagent to each well. The plate was covered and incubated for a total of ten minutes before reading the luminescence with a Flexstation 3 plate reader (All wavelengths mode, 500 ms integration time), graphed using Graphpad Prism 7 (Sigmoidal 4PL mode), and an  $\text{IC}_{50}$  value was calculated.

For the full GTPase inhibition curve for **L2** shown in Figure 4, the above procedure was followed with the change that the incubation with KRas occurred over or four hours rather than two.

## References

- (1) Sano, T.; Cantor, C. R. *Proc. Natl. Acad. Sci.* **1990**, 87 (1), 142.
- (2) *The extinction coefficients for the proteins, peptides, and epitopes were calculated using the peptide properties calculator found at <http://biotools.nubic.northwestern.edu/proteincalc.html>.*
- (3) *The extinction coefficients for the ssDNA strands were calculated using the IDT oligo analyzer at <https://www.idtdna.com/calc/analyzer>.*
- (4) Boriack-Sjodin, P. A.; Margarit, S. M.; Bar-Sagi, D.; Kuriyan, J. *Nature* **1998**, 394

(6691), 337.

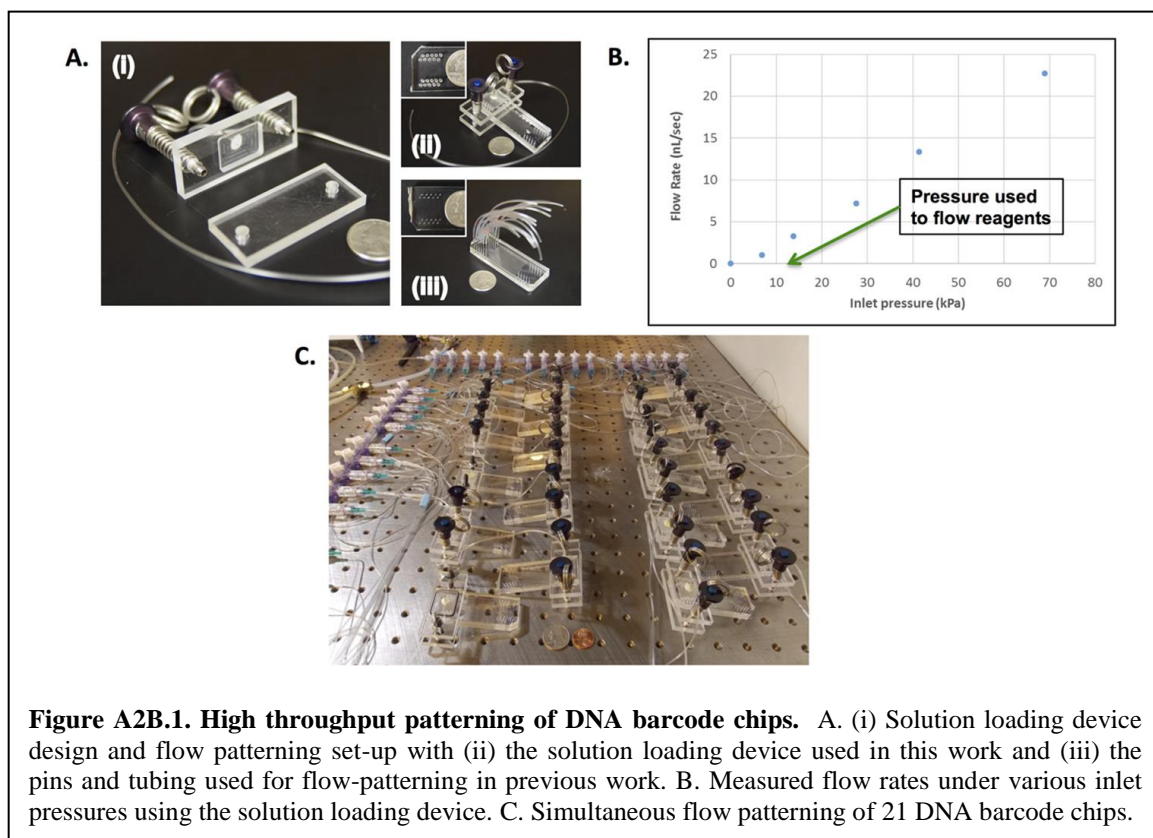
## Appendix 2B

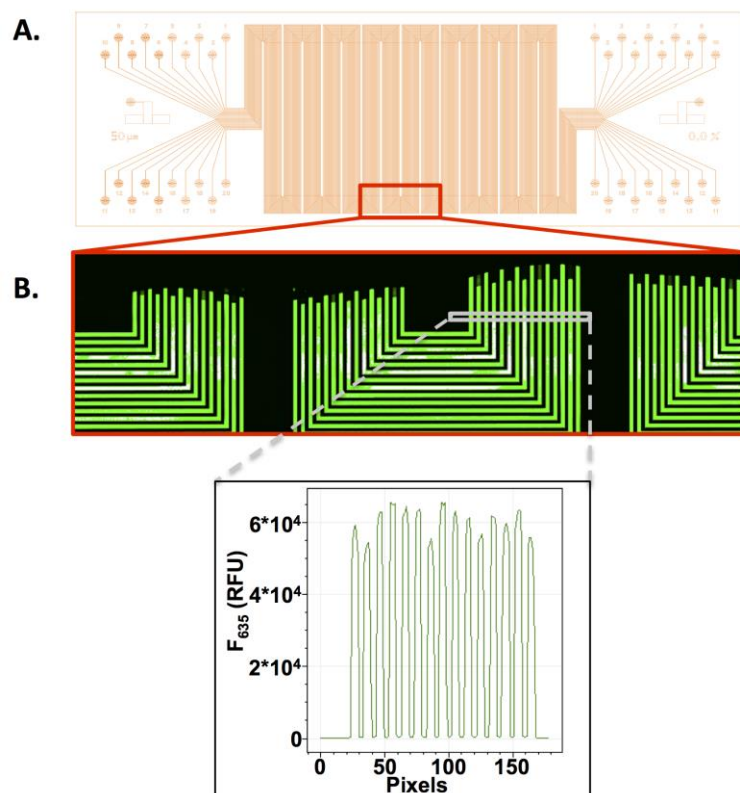
### Supplemental Figures

#### **Table of Contents**

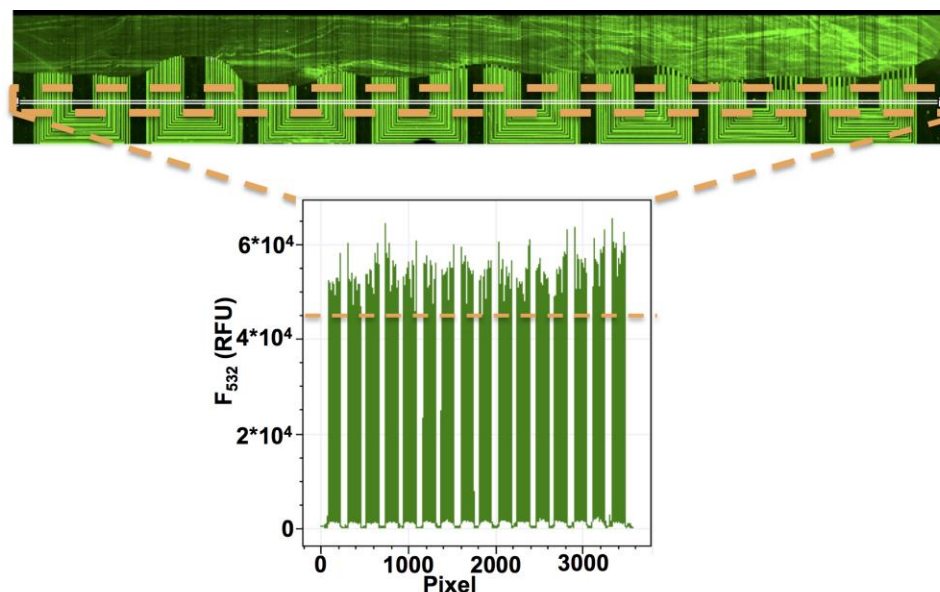
Figure A2B.1	67
Figure A2B.2	68
Figure A2B.3	69
Figure A2B.4	69
Figure A2B.5	70
Figure A2B.6	70
Figure A2B.7	70
Figure A2B.8	71
Figure A2B.9	72
Figure A2B.10	73
Figure A2B.11	74
Figure A2B.12	75
Figure A2B.13	76
Figure A2B.14	77
Figure A2B.15	78
Figure A2B.16	79
Figure A2B.17	80
Figure A2B.18	81
Figure A2B.19	82
Figure A2B.20	83
Figure A2B.21	84
Figure A2B.22	84
Figure A2B.23	85
Figure A2B.24	86
Figure A2B.25	87
Figure A2B.26	87
Figure A2B.27	88
Figure A2B.28	88



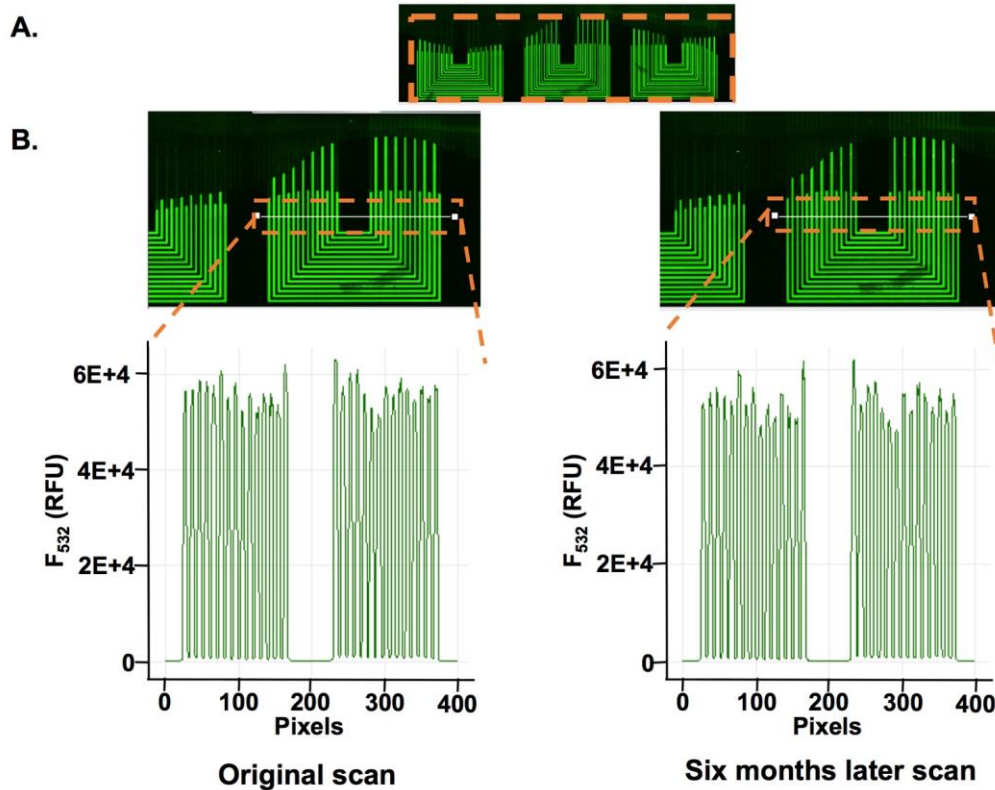




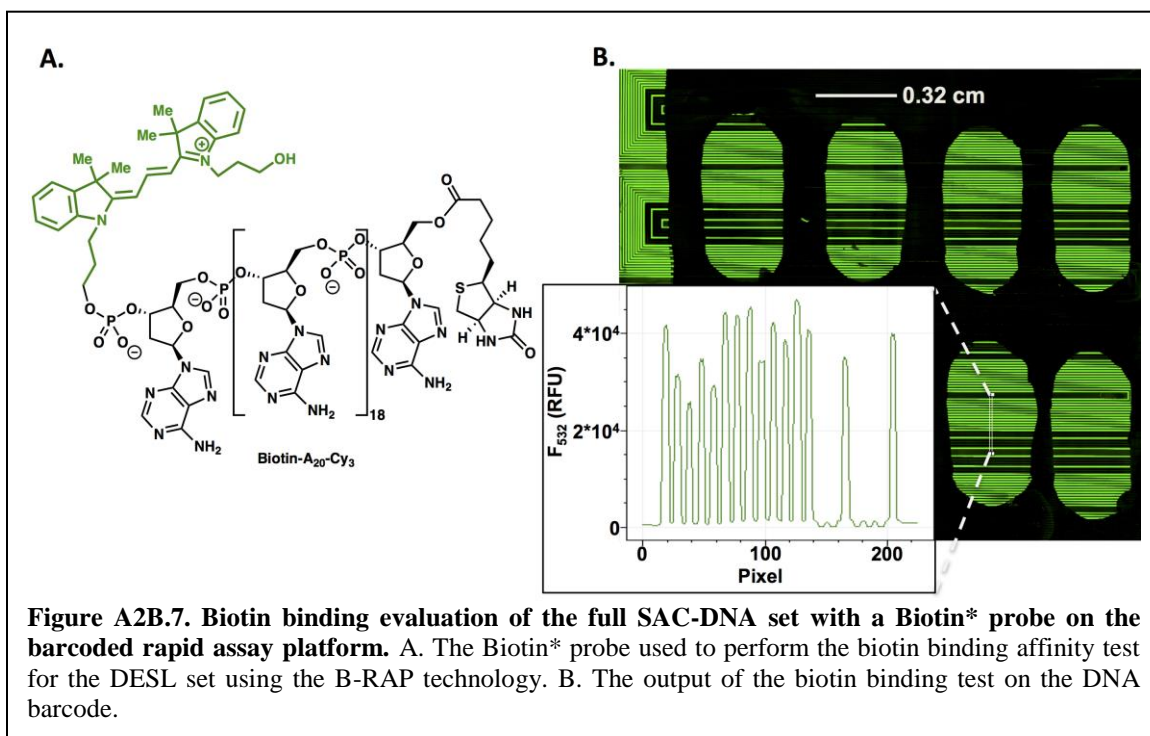
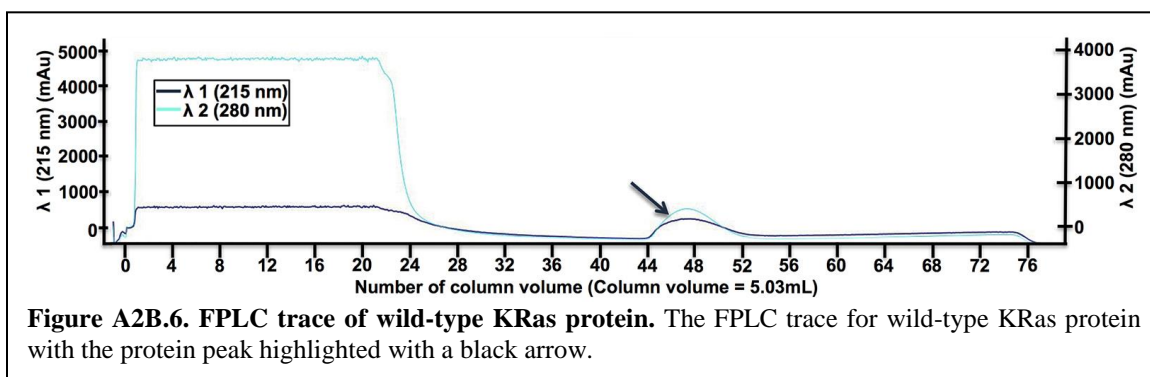
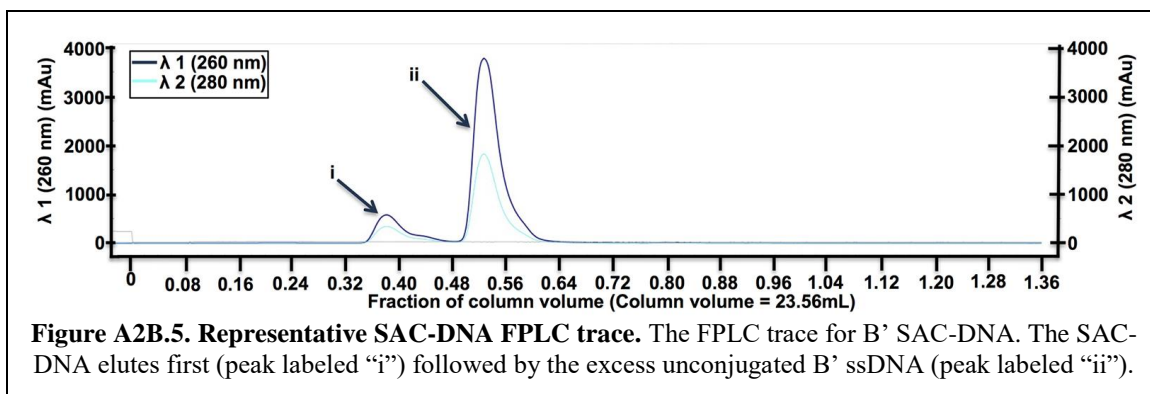
**Figure A2B.2. DNA barcode chip layout and validation.** A. The 50  $\mu\text{m}$  barcode chip layout, which encompasses the entire length of a 3" microscope slide. Input and outputs of the serpentine microchannels are at the right and left sides. B. An image and digitized line scan of a barcoded region of the slide. The white box in the image outlines a region that is approximately 50  $\mu\text{m}$  x 2 mm wide.

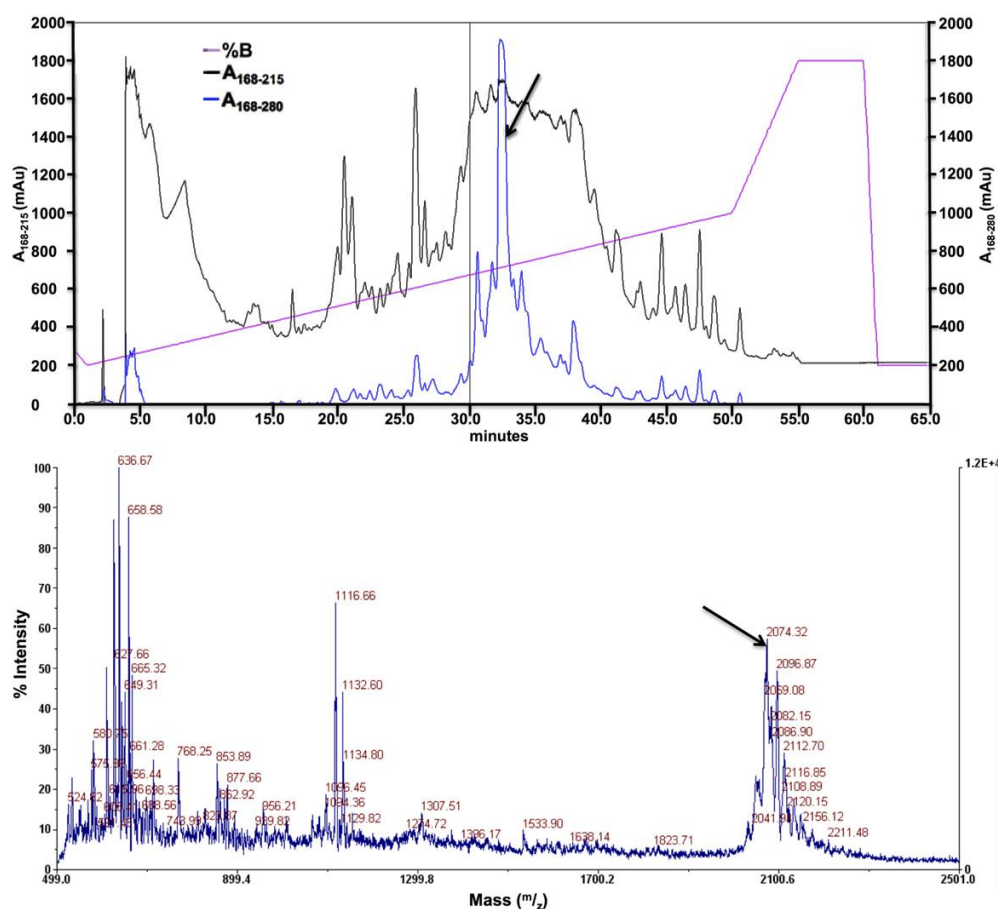


**Figure A2B.3. Comparison of barcode quality across an entire microchip.** Measuring the barcode quality across the entire barcode region (about 3.8 cm) reveals that all of the barcode lanes meet the minimum  $F_{532}$  requirement for good levels of ssDNA patterning as highlighted by the dashed line on the graph.

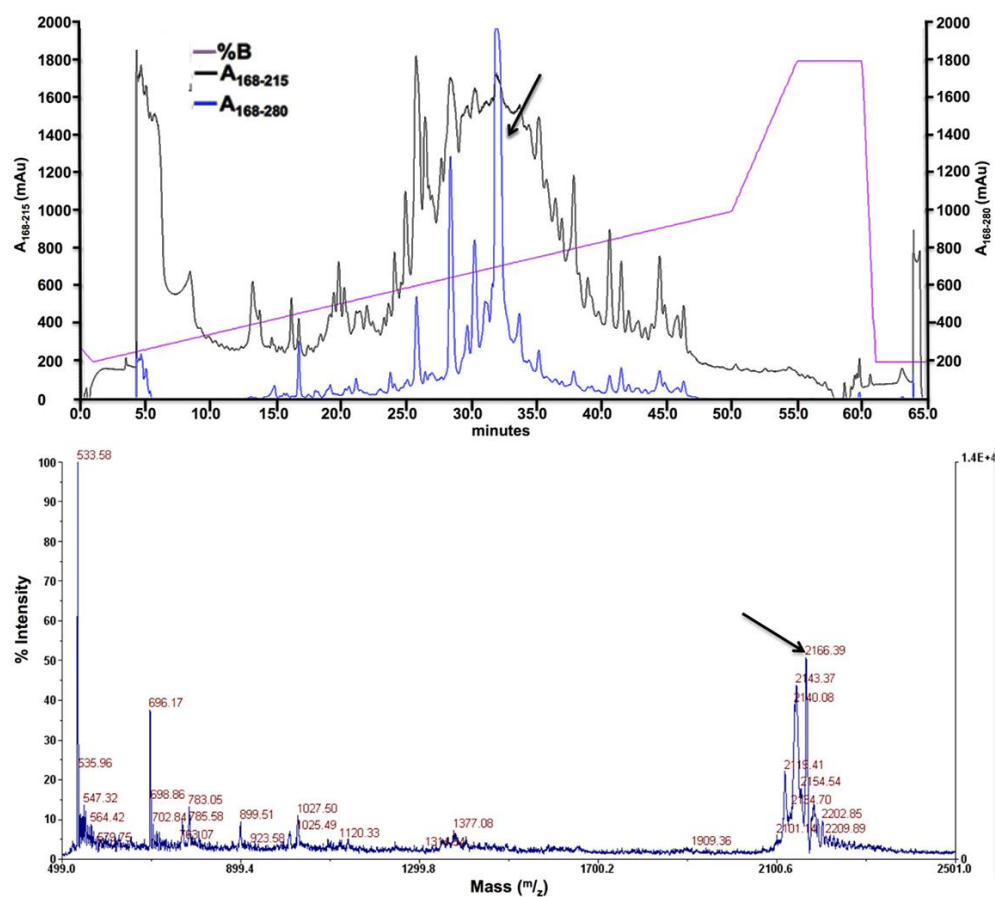


**Figure A2B.4. Probing barcode stability.** A. The portion of the validation region used to generate Table A2C.4. B. A comparison of  $F_{532}$  output for two DNA barcode sets in the validation region with six months between each scan.

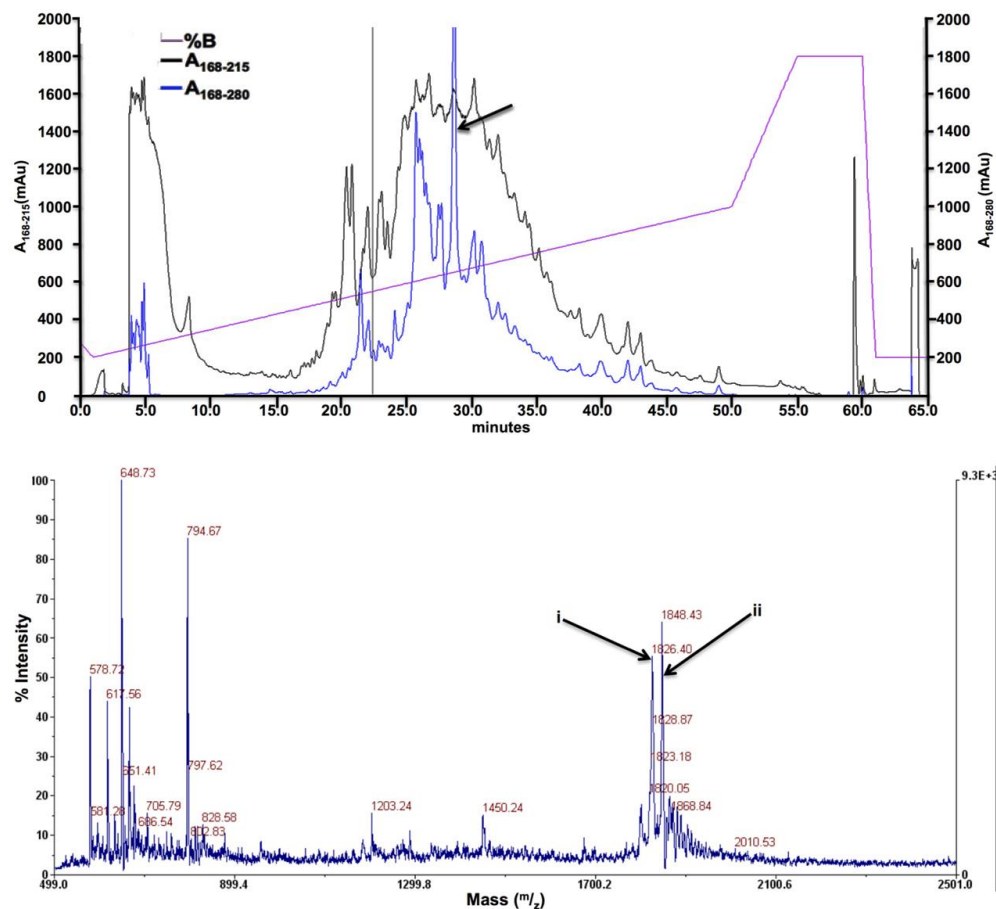




**Figure A2B.8.** The characterization data for SynEp1. SynEp1 FDEYD[P→<sup>4-N</sup><sup>3</sup>P]TIEDSY-PEG<sub>4</sub>-Biotin: C<sub>90</sub>H<sub>128</sub>N<sub>20</sub>O<sub>33</sub>S. (Top) The semi-preparative HPLC trace for SynEp1 with the peak containing the correct mass indicated by an arrow. (Bottom) The MALDI-TOF spectra for SynEp1 with the peak indicated with an arrow corresponding to  $[M+Na]^+$ , expected  $[M+Na]=2071.188$  amu.

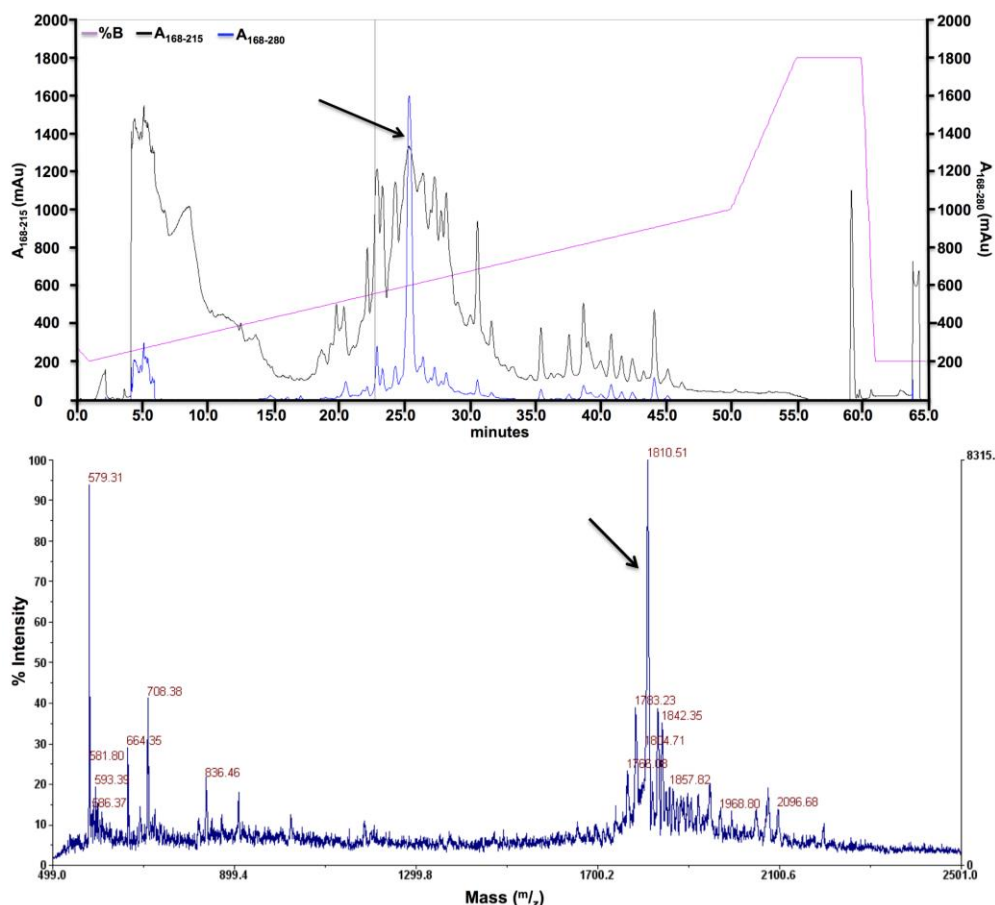


**Figure A2B.9. The characterization data for scrambled SynEp1.** Scrambled SynEp1 [ $P \rightarrow 4$ - $N^3$ P]EYDSIDDEFYT-PEG<sub>4</sub>-Biotin: C<sub>90</sub>H<sub>128</sub>N<sub>20</sub>O<sub>33</sub>S. (Top) The semi-preparative HPLC trace for scrambled SynEp1 with the peak containing the correct mass indicated with an arrow. (Bottom) The MADLI-TOF spectra for scrambled SynEp1. The peak indicated with an arrow corresponds to  $[M+2\text{Pip}-3\text{ OH}+H]$ , expected  $[M+2\text{Pip}-3\text{ OH}+H]=2167.030$  amu, which results from multiple piperidine and aspartamide adducts forming on the oligopeptide.



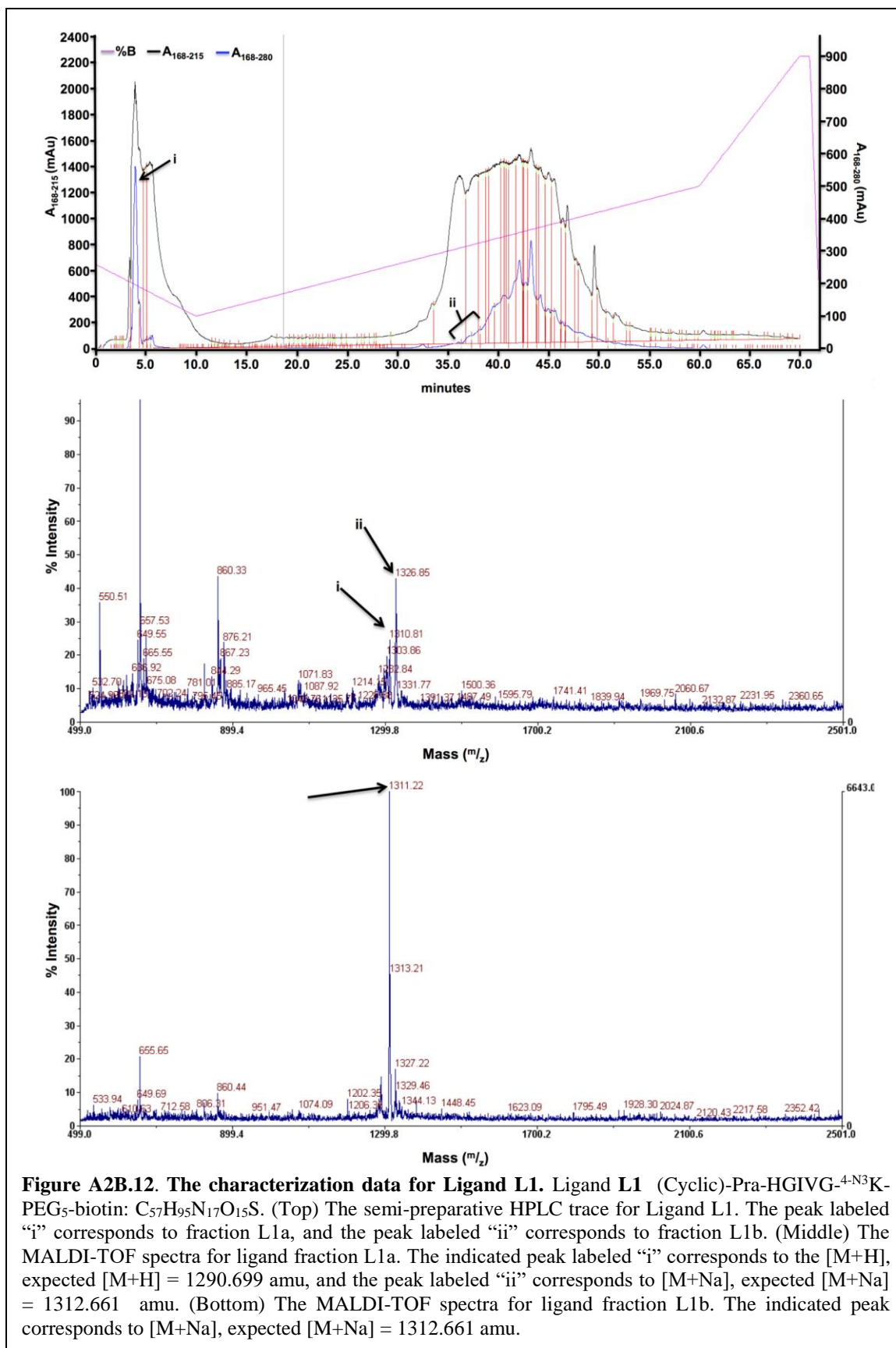
**Figure A2B.10.** The characterization data for **SynEp2**. **SynEp2** [ $I \rightarrow 4\text{-N}_3$ K]LDTAGQEEYS-PEG<sub>5</sub>-Biotin:  $C_{77}H_{124}N_{20}O_{29}S$ . (Top) The semi-preparative HPLC trace for **SynEp2** with the peak containing the correct mass indicated by an arrow. (Bottom) The MALDI-TOF spectra for **SynEp2**. The peak labeled “i” corresponds to  $[M+H]$ , expected  $[M+H]=1825.864$  amu, and the peak labeled “ii” corresponds to  $[M+Na]$ , expected  $[M+Na]=1847.846$  amu.

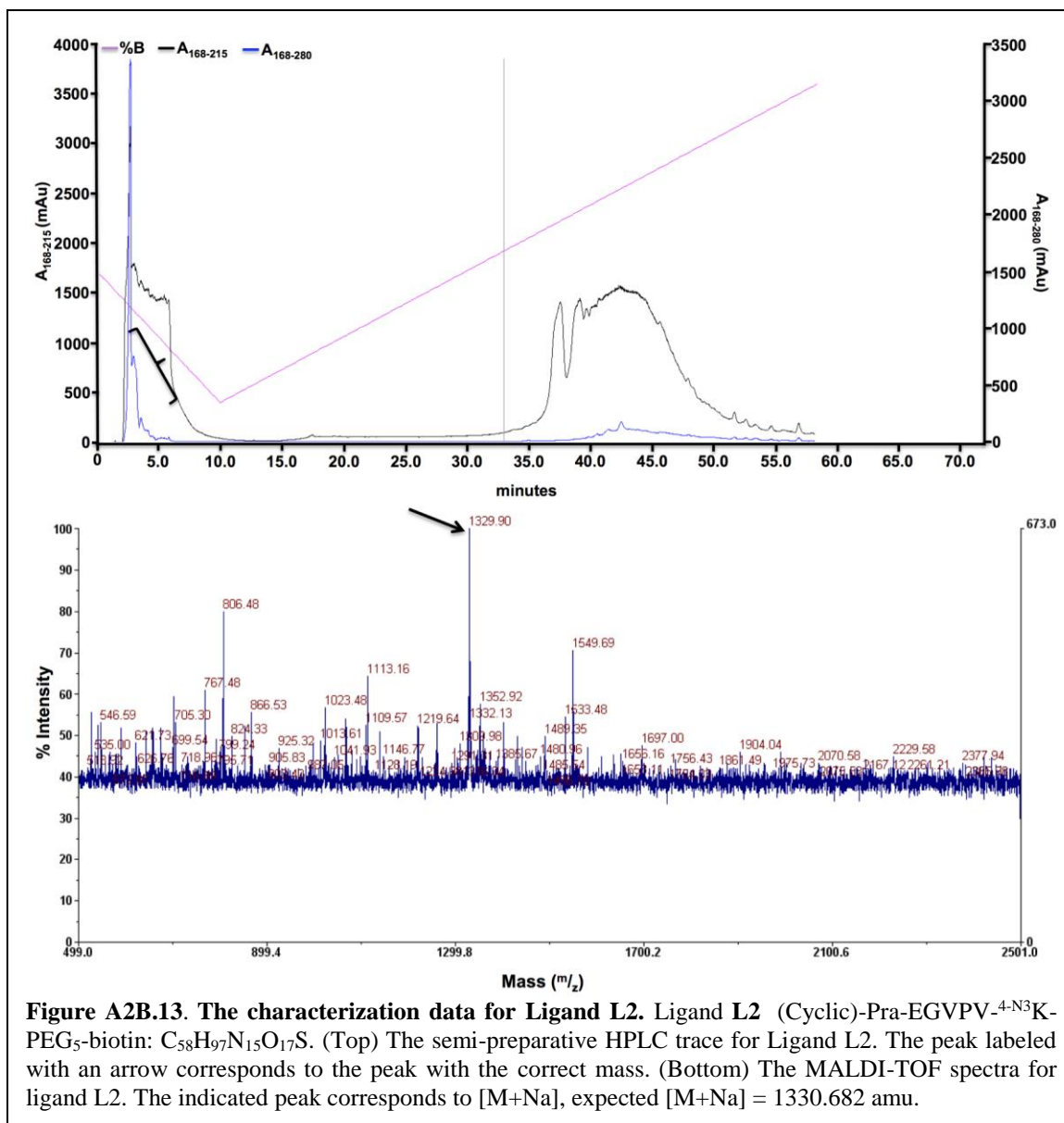




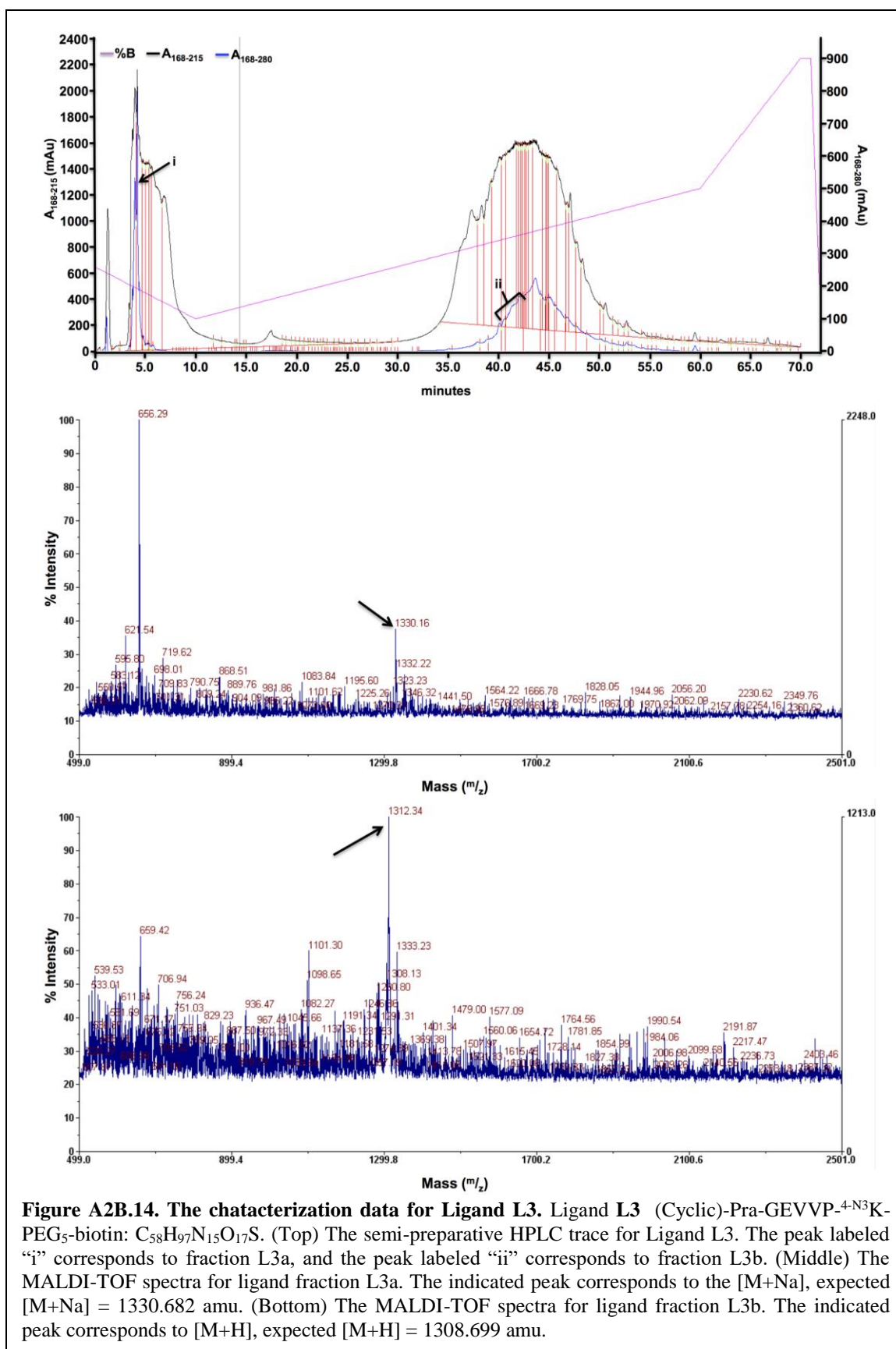
**Figure A2B.11. The characterization data for scrambled SynEp2. Scrambled SynEp2 [ $I \rightarrow 4$ - $N^3K$ ]LSTGEYDAQE-PEG<sub>5</sub>-Biotin:  $C_{77}H_{124}N_{20}O_{29}S$ .** (Top) The semi-preparative HPLC trace for scrambled SynEp2 with the peak containing the correct mass indicated by an arrow. (Bottom) The MALDI-TOF spectra for scrambled SynEp2 with the indicated peak corresponding to the epitope with a single asparamide formation  $[M-OH+H]$ , expected  $[M-OH+H]=1809.99$  amu.

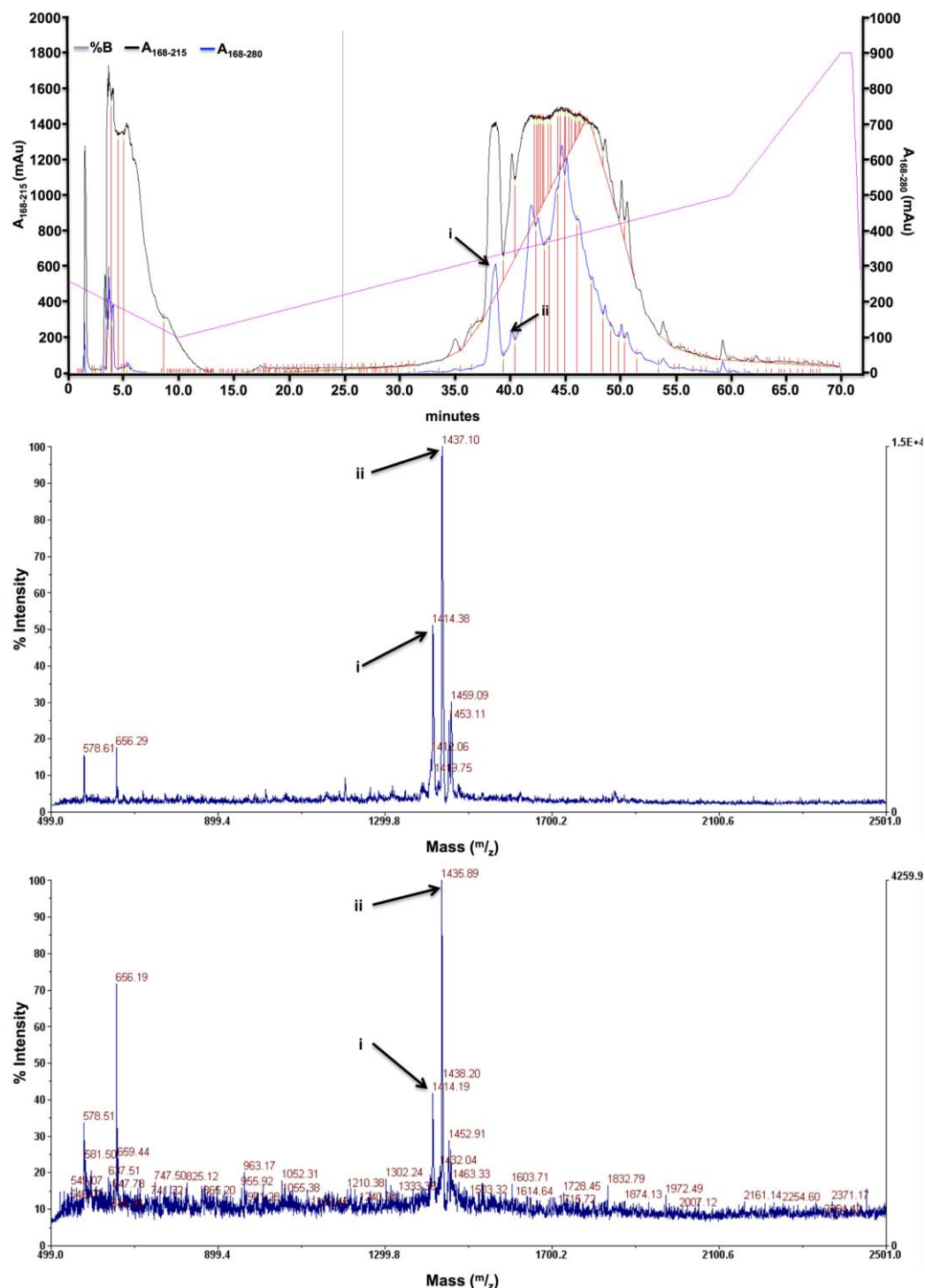




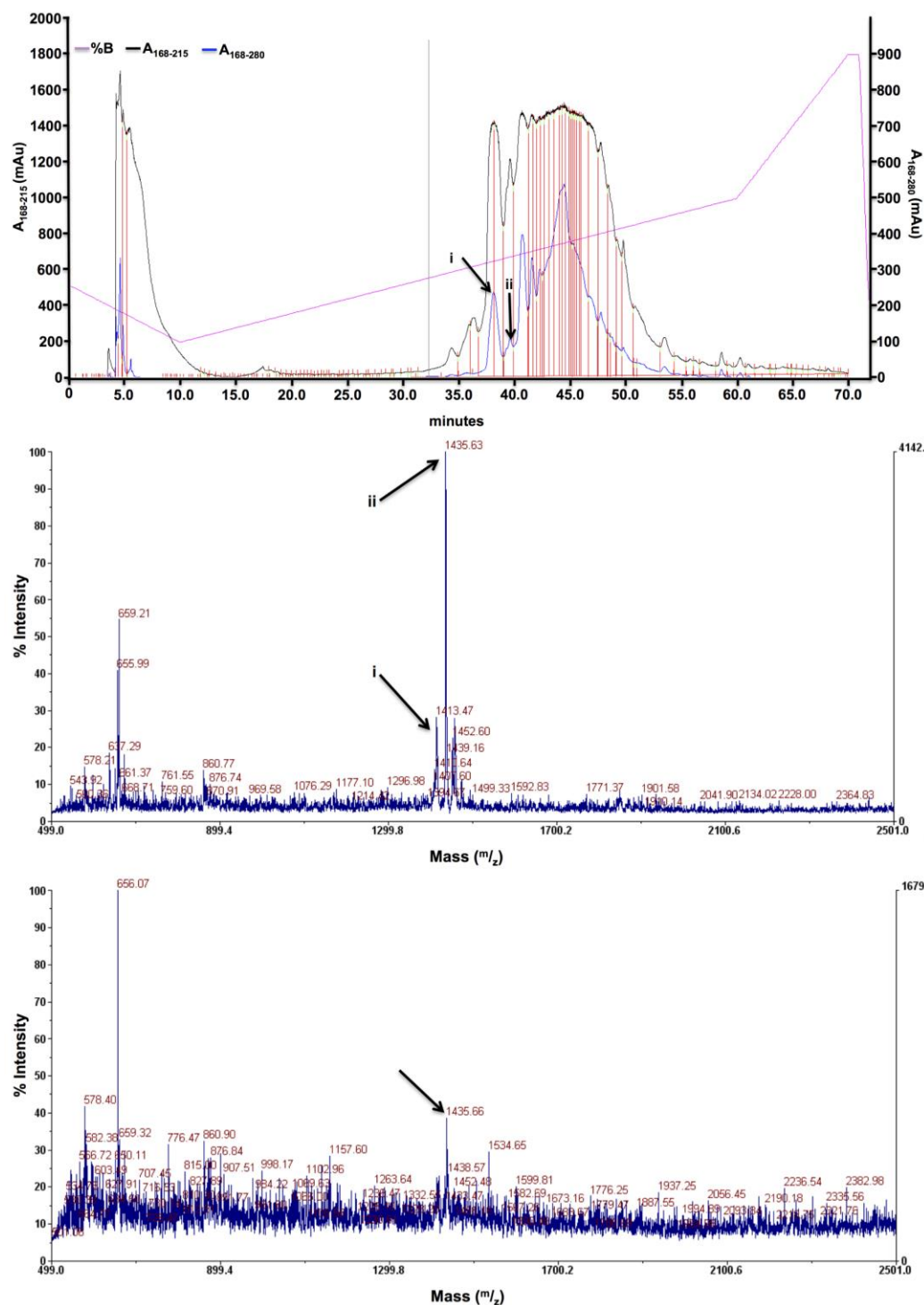


**Figure A2B.13. The characterization data for Ligand L2.** Ligand L2 (Cyclic)-Pra-EGVPV-<sup>4-N<sub>3</sub></sup>K-PEG<sub>5</sub>-biotin: C<sub>58</sub>H<sub>97</sub>N<sub>15</sub>O<sub>17</sub>S. (Top) The semi-preparative HPLC trace for Ligand L2. The peak labeled with an arrow corresponds to the peak with the correct mass. (Bottom) The MALDI-TOF spectra for ligand L2. The indicated peak corresponds to [M+Na], expected [M+Na] = 1330.682 amu.

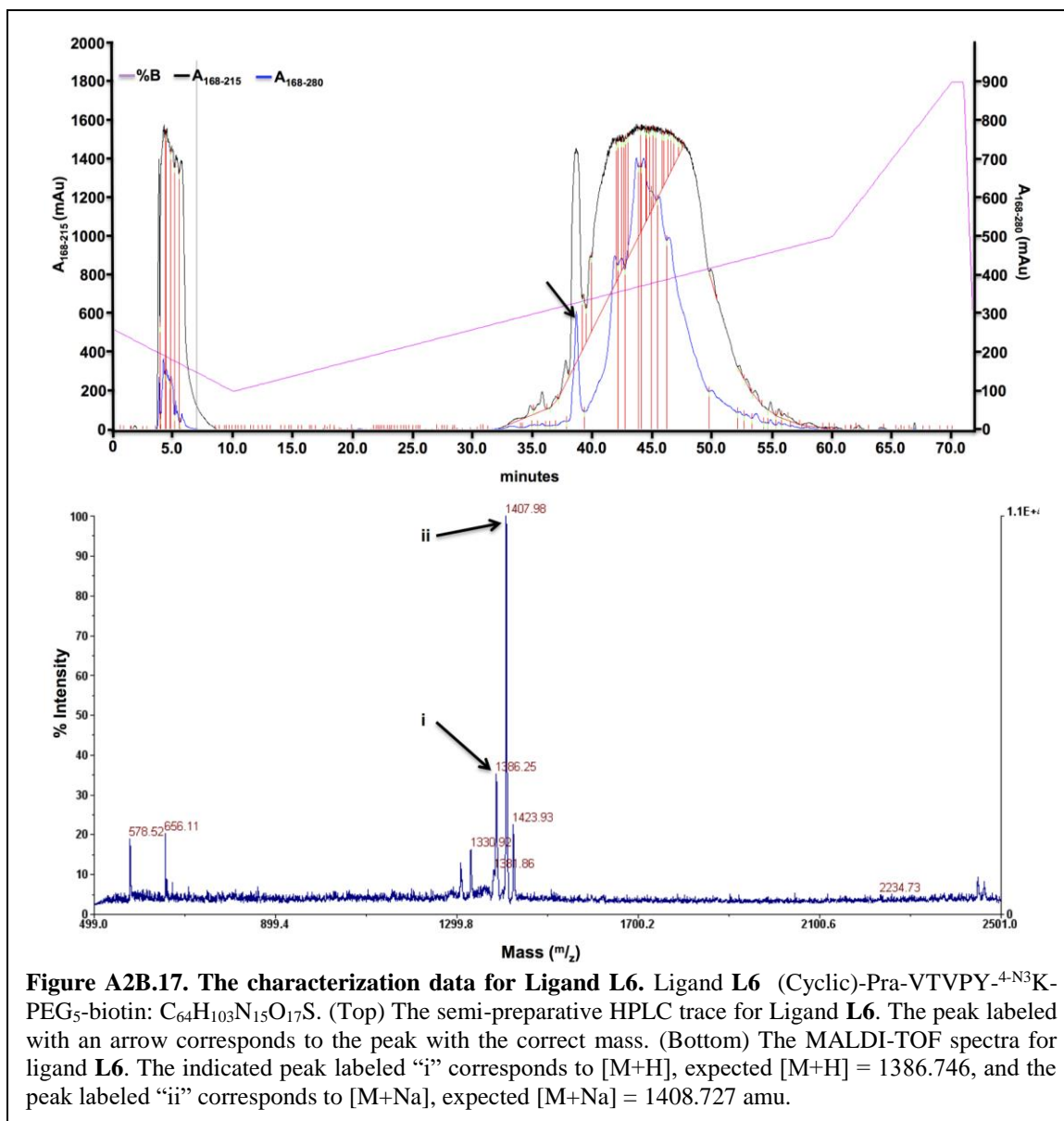




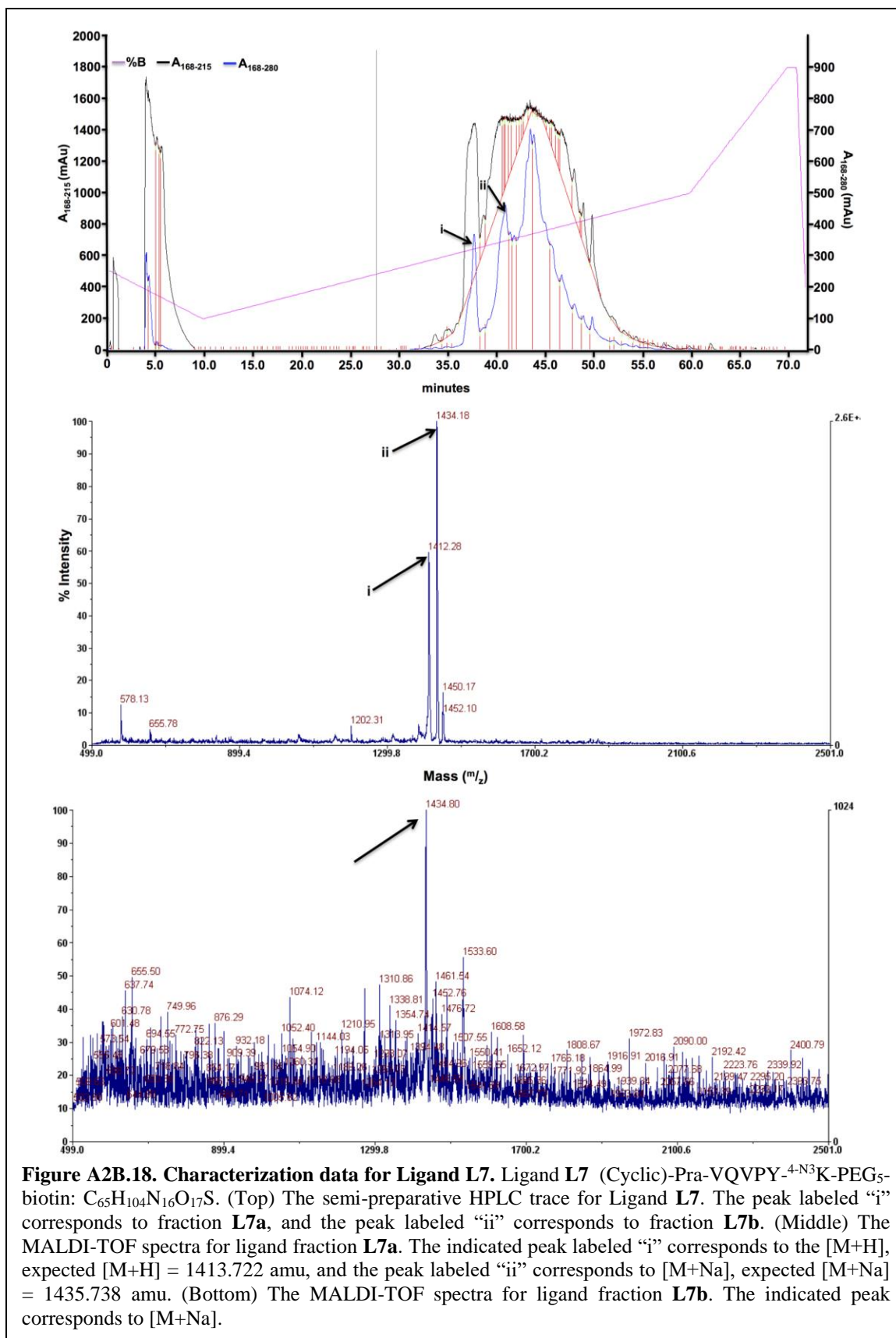
**Figure A2B.15. The characterization data for Ligand L4.** Ligand L4 (Cyclic)-Pra-VEVPY-<sup>4-N</sup><sub>3</sub>K-PEG<sub>5</sub>-biotin: C<sub>65</sub>H<sub>103</sub>N<sub>15</sub>O<sub>18</sub>S. (Top) The semi-preparative HPLC trace for Ligand L4. The peak labeled “i” corresponds to fraction L4a, and the peak labeled “ii” corresponds to fraction L4b. (Middle) The MALDI-TOF spectra for ligand fraction L4a. The indicated peak labeled “i” corresponds to the [M+H]<sup>+</sup>, expected [M+H] = 1414.740 amu, and the peak labeled “ii” corresponds to [M+Na]<sup>+</sup>, expected [M+Na] = 1436.722 amu. (Bottom) The MALDI-TOF spectra for ligand fraction L4b. The indicated peak labeled “i” corresponds to the [M+H]<sup>+</sup>, and the peak labeled “ii” corresponds to [M+Na]<sup>+</sup>.

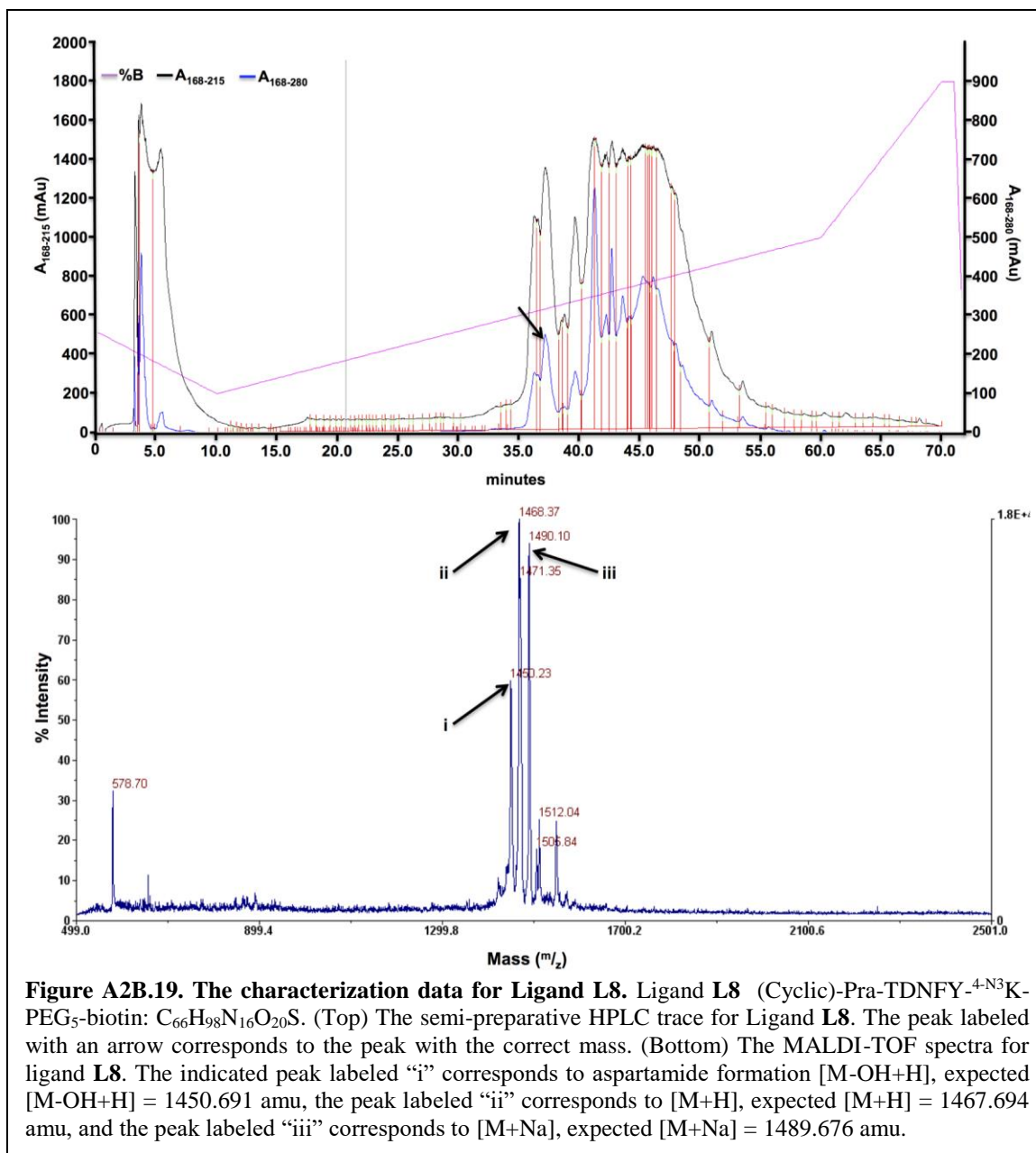


**Figure A2B.16. Characterization data for Ligand L5.** Ligand L5 (Cyclic)-Pra-VEVYP-<sup>4-N</sup><sup>3</sup>K-PEG<sub>5</sub>-biotin: C<sub>65</sub>H<sub>103</sub>N<sub>15</sub>O<sub>18</sub>S. (Top) The semi-preparative HPLC trace for Ligand L5. The peak labeled “i” corresponds to fraction L5a, and the peak labeled “ii” corresponds to fraction L5b. (Middle) The MALDI-TOF spectra for ligand fraction L5a. The indicated peak labeled “i” corresponds to the [M+H]<sup>+</sup>, expected [M+H]<sup>+</sup> = 1414.740 amu, and the peak labeled “ii” corresponds to [M+Na]<sup>+</sup>, expected [M+Na]<sup>+</sup> = 1436.722 amu. (Bottom) The MALDI-TOF spectra for ligand fraction L4b. The indicated peak corresponds to [M+Na]<sup>+</sup>.

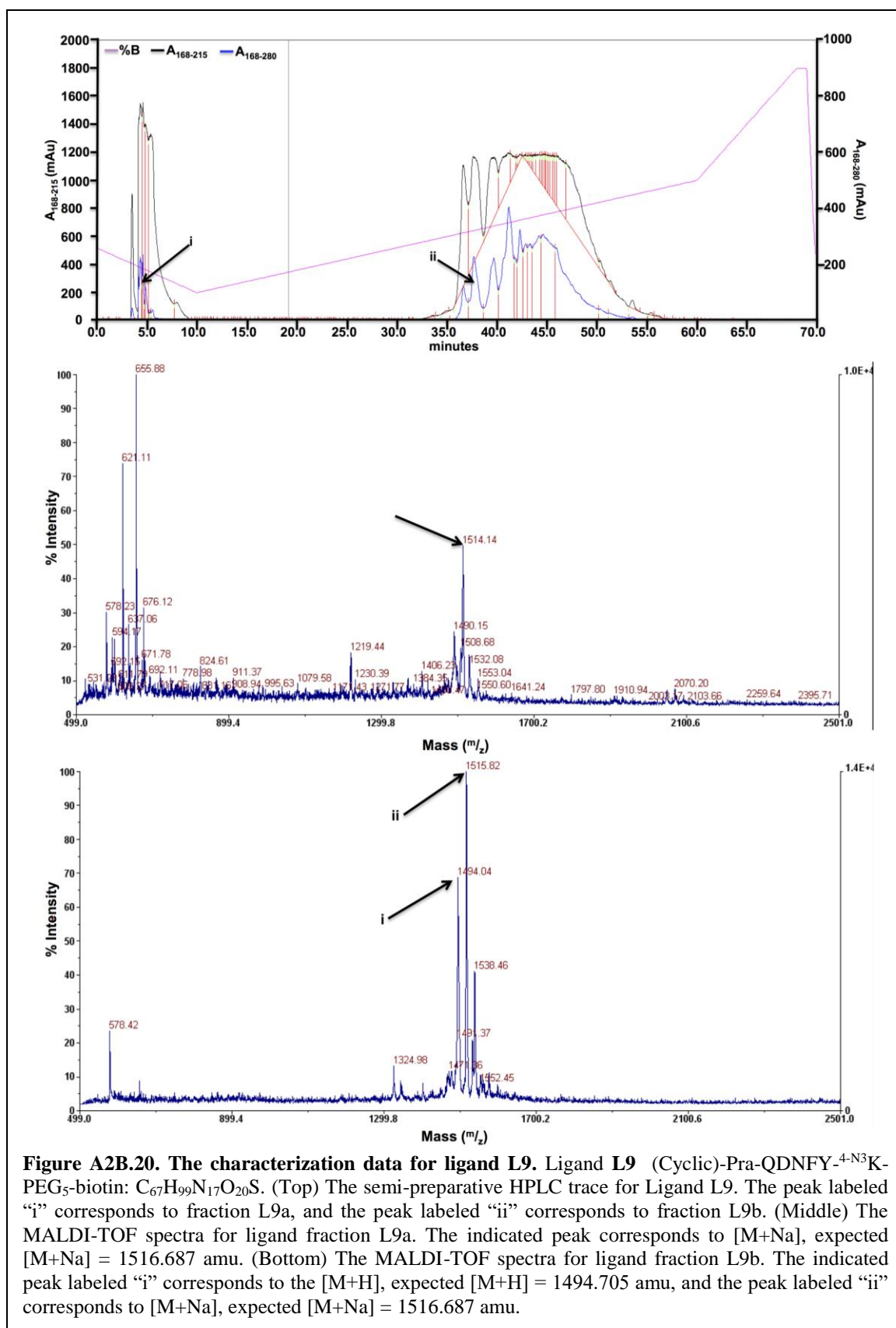


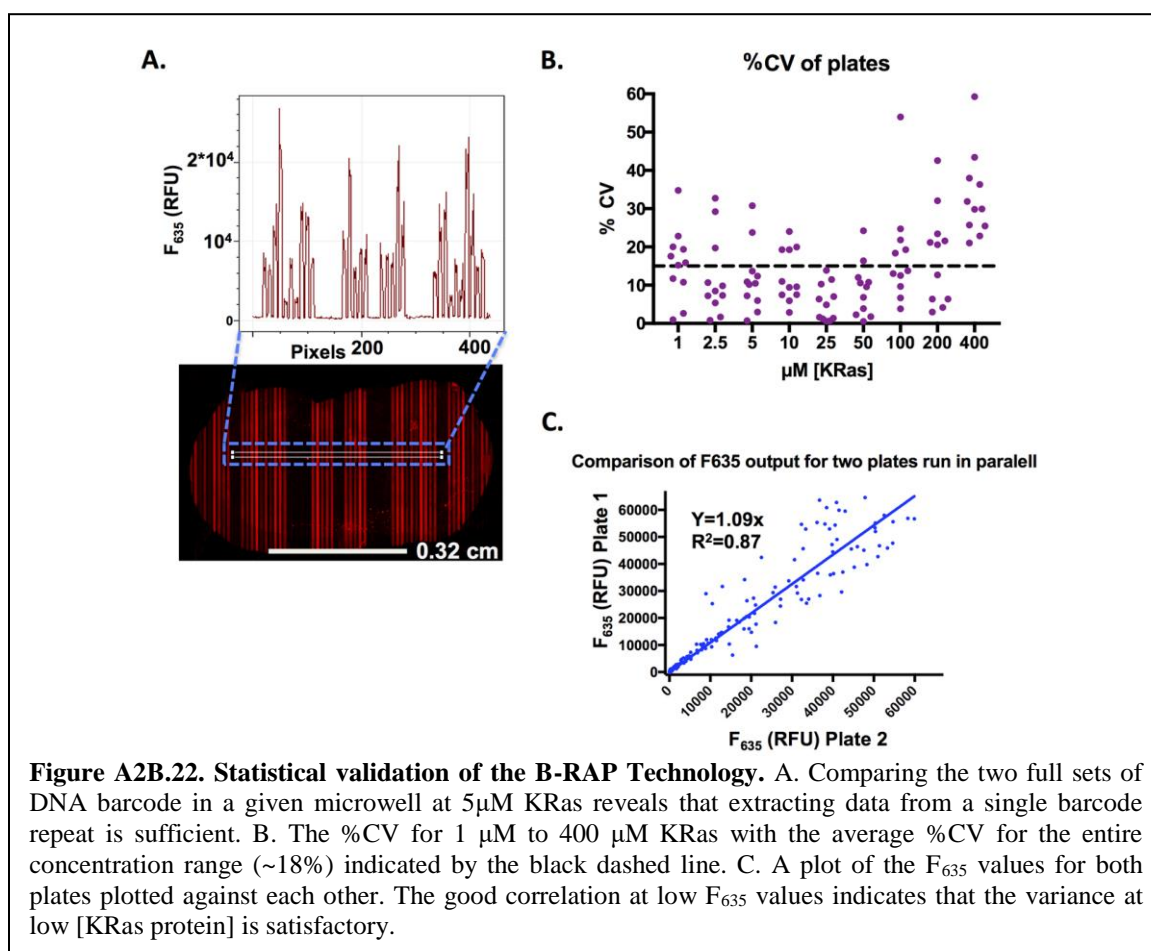
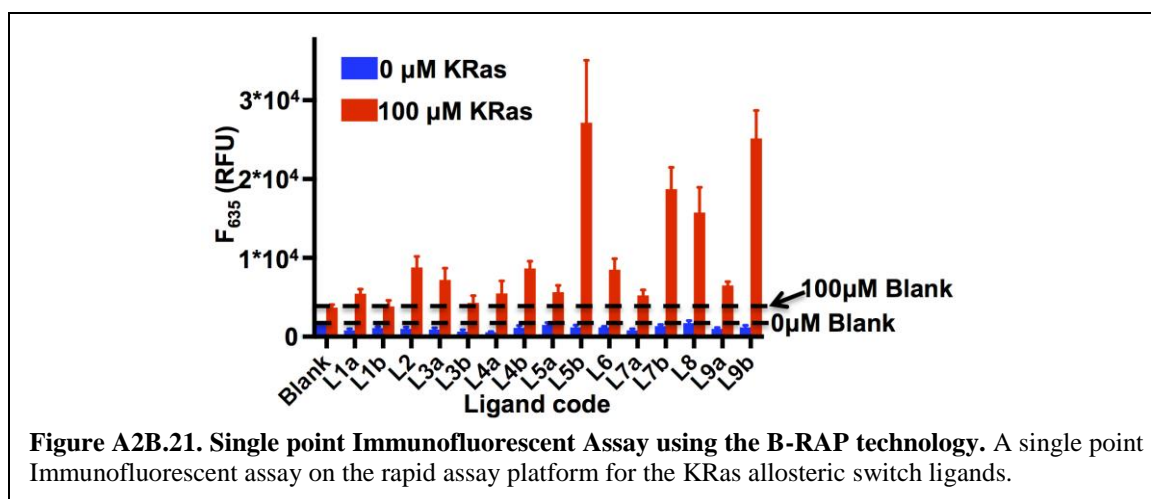


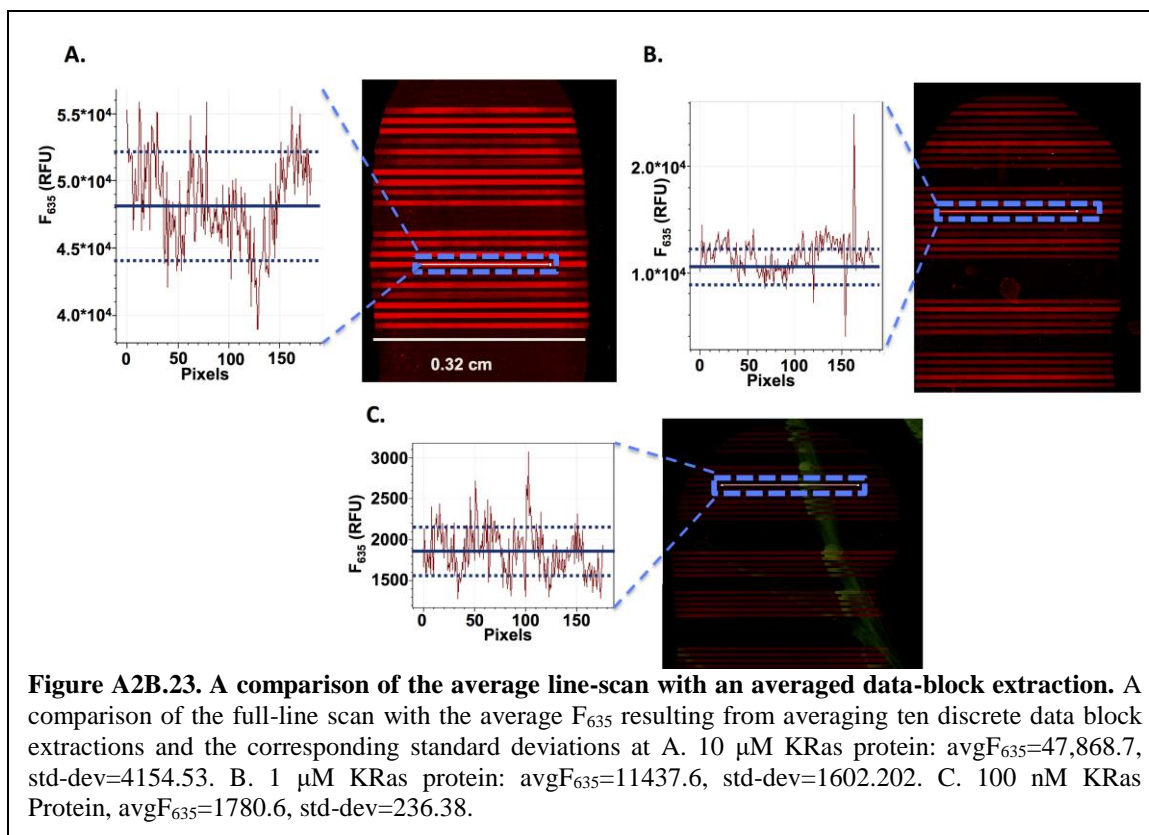


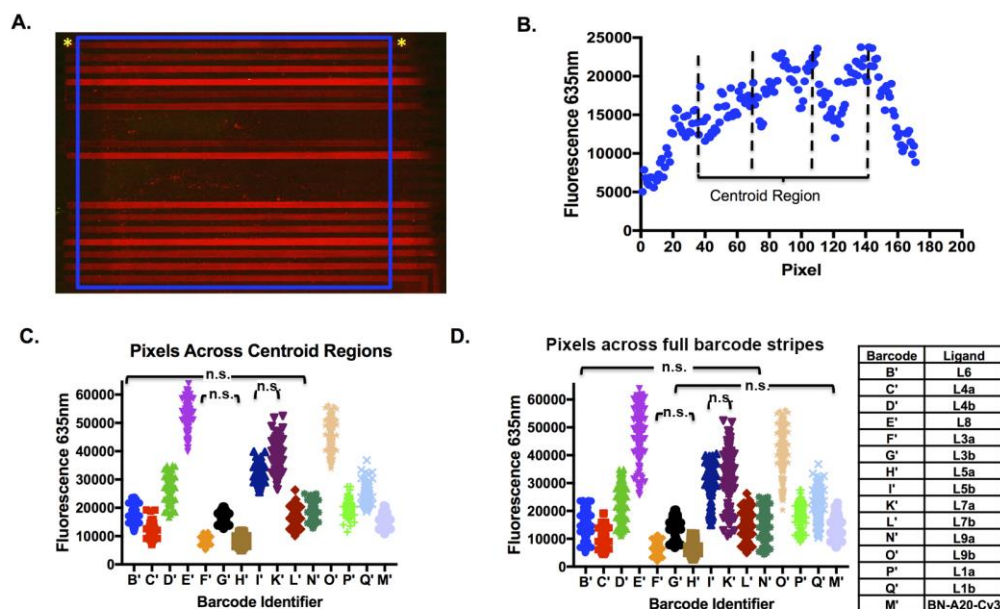




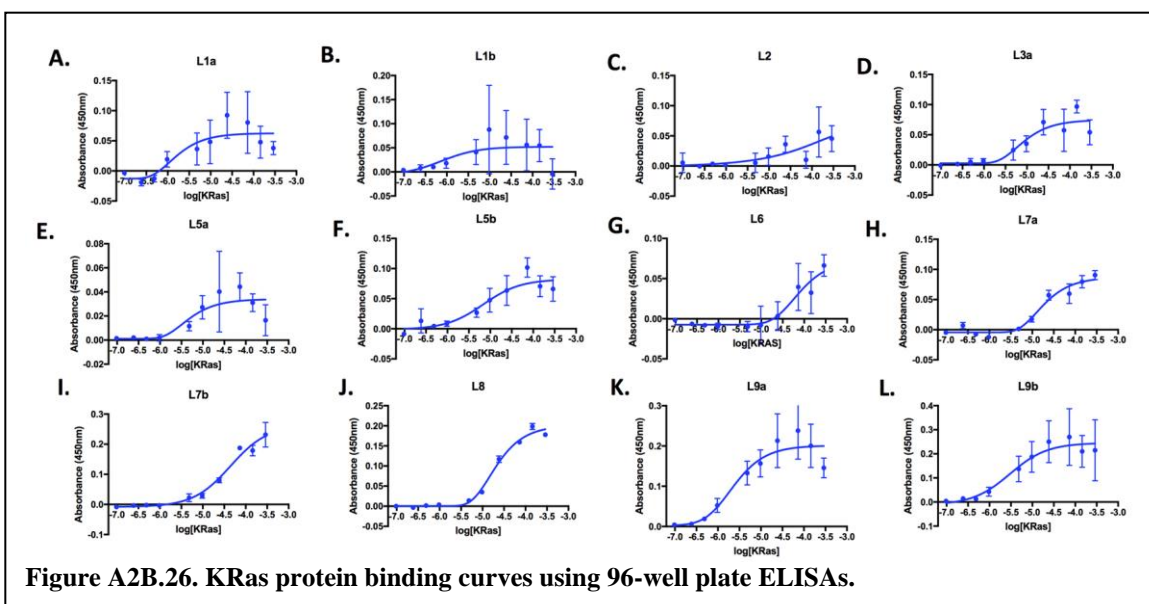
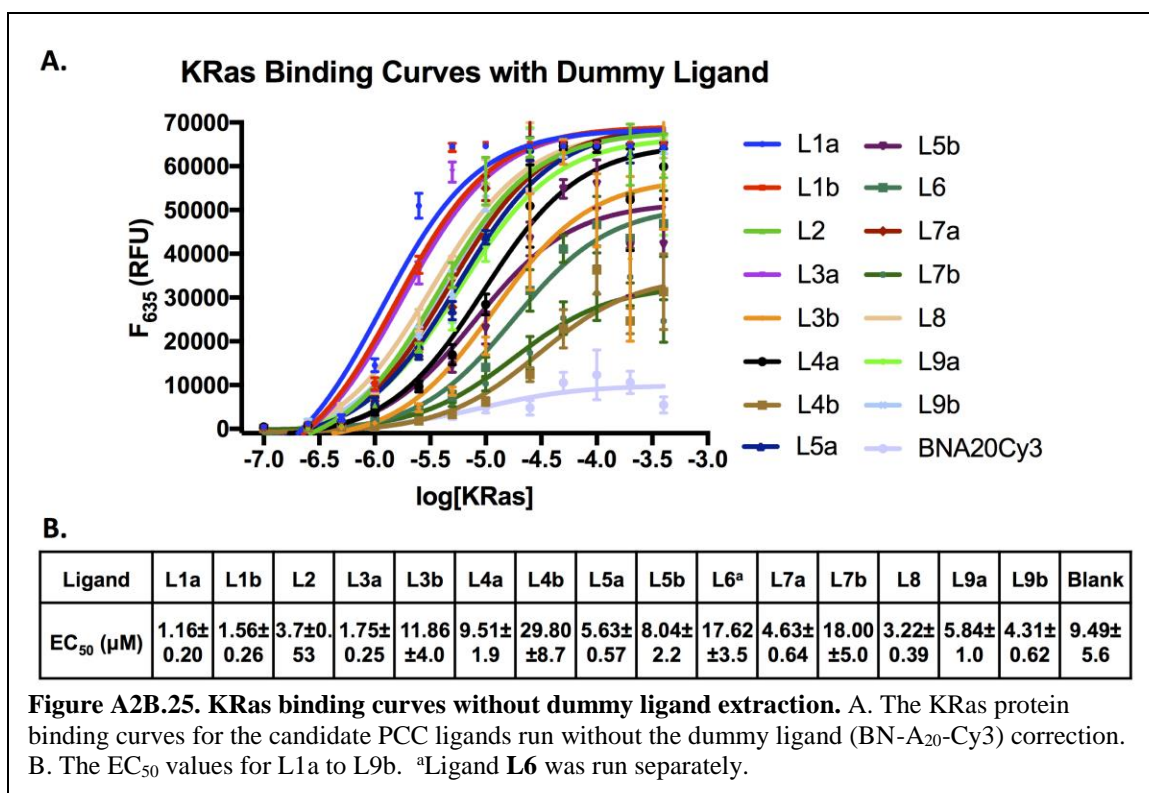


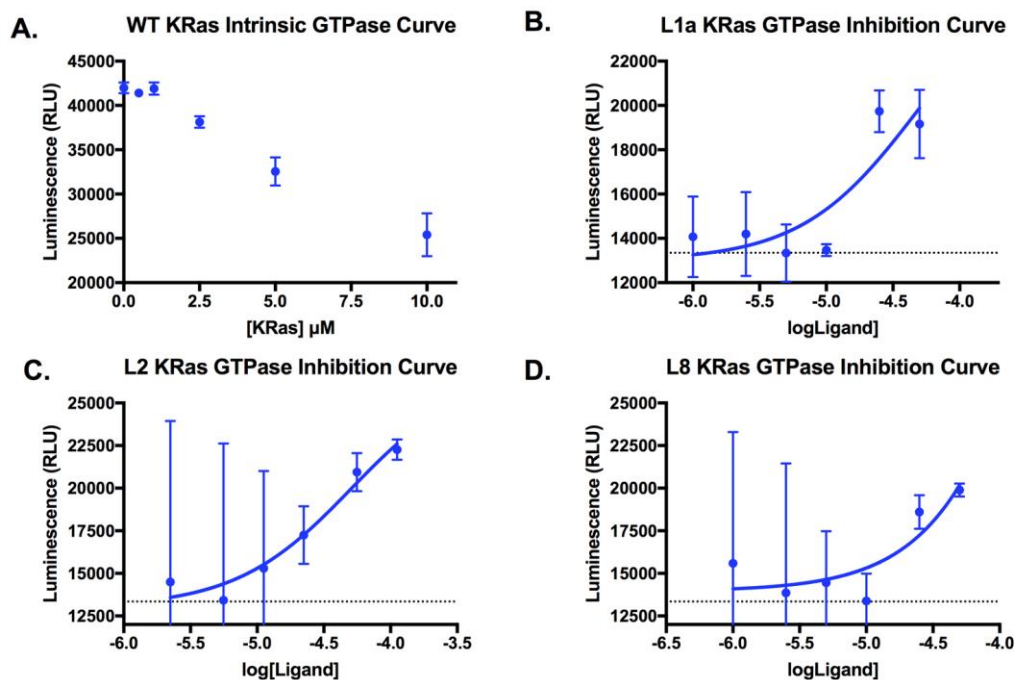




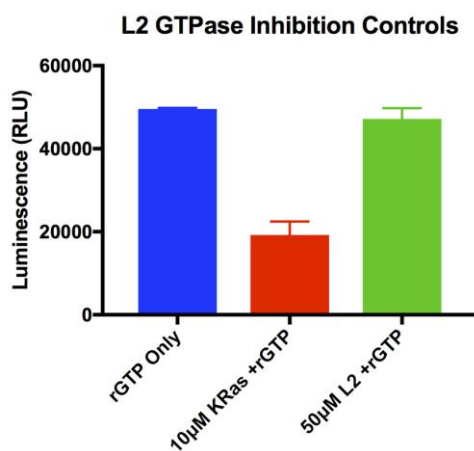


**Figure A2B.24. Pixel extraction of full barcode lanes in a given barcode.** A. The region of the barcode lanes extracted with the B' lane graphed in a) set off by yellow asterisks. B. A close up of the scatterplot of barcode lane B' to highlight where the centroid region was defined and extracted (pixel 40-142) with the region divided into thirds for the left, middle, and right centroid respectively C. A scatterplot of the F635 extracted for individual pixels along the entire length of a barcode in the 10  $\mu$ M well of plate #1 with lanes that were indistinguishable by a two-tailed student T test indicated by brackets, and an accompanying key for parts c) and d). D. A scatterplot of the F635 extracted for individual pixels in the centroid region of each barcode with lanes that were indistinguishable by a two-tailed student T-test (Table S5) indicated by brackets.





**Figure A2B.27. The wild-type KRas protein GTPase activity curves.** A. The WT Kras standard curve generated with a one hour incubation with 5  $\mu$ M rGTP. The inhibition curves generated for the best three ligand fractions with a two hour rGTP incubation period in B., C., and D. The dotted line represents the luminescence measurement (~13k RLU) when 10  $\mu$ M KRas was incubated with 5  $\mu$ M rGTP for two hours without ligand.



**Figure A2B.28. Controls for the L2 GTPase inhibition curve.** Controls run for the L2 inhibition curve. It is apparent that ligand L2 lacks GTPase catalytic activity as less than 5% of the GTP was hydrolyzed when placed with just 50  $\mu$ M of L2.

## **Appendix 2C**

### **Supplemental Tables**

#### **Table of Contents**

Table A2C.1	90
Table A2C.2	91
Table A2C.3	92
Table A2C.4	93
Table A2C.5	95
Table A2C.6	95
Table A2C.7	96
Table A2C.8	97

<b>DNA Patterning step</b>	<b>Solution Loading Method (new)</b>	<b>Tubing Method (old)</b>
Polylysine loading	12 min	> 24 min
Polylysine flow setting	10 min	20 min
DNA tubing preparation	n.a.	30 min
DNA loading	22 min	50 min

**Table A2C.1. A comparison of the various steps associated with the Solution Loading Method (reported here) and the previously reported Tubing Method for preparing microfluidic flow-patterned ssDNA barcodes.** By using the solution loading method, 20-25 devices are handled within one hour as opposed to 8-10 devices with individual tubing for each amine DNA.



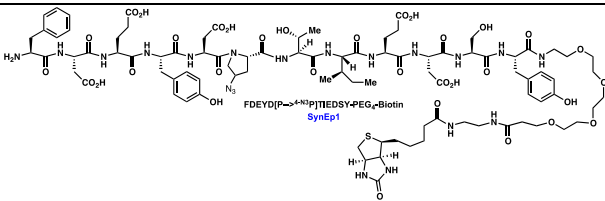
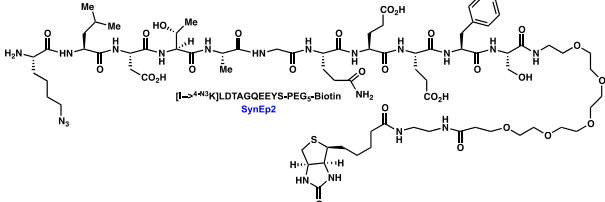
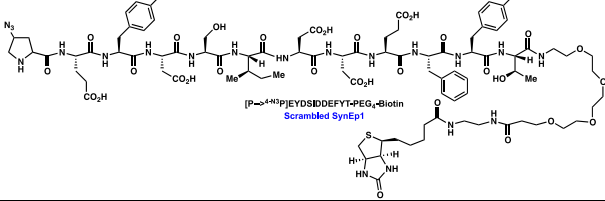
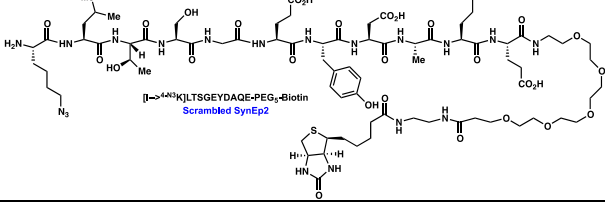
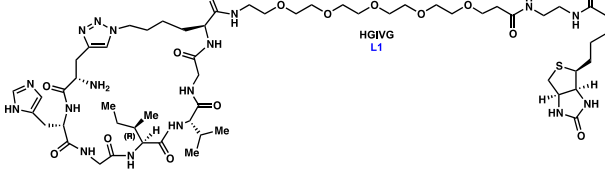
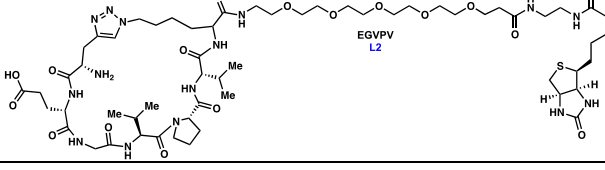
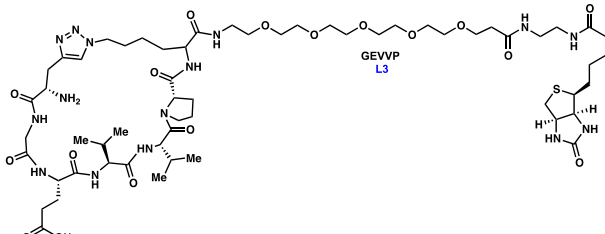
DNA i.d.	Sequence
<b>B</b>	5'-NH2-C6-AAA AAA AAA AAA AGC CTC ATT GAA TCA TGC CTA-3'
<b>B'</b>	5'-NH2-C6-AAA AAA AAA ATA GGC ATG ATT CAA TGA GGC-3'
<b>C</b>	5'-NH2-C6-AAA AAA AAA AAA AGC ACT CGT CTA CTA TCG CTA-3'
<b>C'</b>	5'-NH2-C6-AAA AAA AAA ATA GCG ATA GTA GAC GAG TGC-3'
<b>D</b>	5'-NH2-C6-AAA AAA AAA AAA GGT CGA GAT GTC AGA GTA-3'
<b>D'</b>	5'-NH2-C6-AAA AAA AAA ATA CTC TGA CAT CTC GAC CAT-3'
<b>E</b>	5'-NH2-C6-AAA AAA AAA AAA AAT GTG AAG TGG CAG TAT CTA-3'
<b>E'</b>	5'-NH2-C6-AAA AAA AAA ATA GAT ACT GCC ACT TCA CAT-3'
<b>F</b>	5'-NH2-C6-AAA AAA AAA AAA AAT CAG GTA AGG TTC ACG GTA-3'
<b>F'</b>	5'-NH2-C6-AAA AAA AAA ATA CCG TGA ACC TTA CCT GAT-3'
<b>G</b>	5'-NH2-C6-AAA AAA AAA AAA AGA GTA GCC TTC CCG AGC ATT-3'
<b>G'</b>	5'-NH2-C6-AAA AAA AAA AAA TGC TCG GGA AGG CTA CTC-3'
<b>H</b>	5'-NH2-C6-AAA AAA AAA AAA AAT TGA CCA AAC TGC GGT GCG-3'
<b>H'</b>	5'-NH2-C6-AAA AAA AAA ACG CAC CGC AGT TTG GTC AAT-3'
<b>I</b>	5'-NH2-C6-AAA AAA AAA AAA ATG CCC TAT TGT TGC GTC GGA-3'
<b>I'</b>	5'-NH2-C6-AAA AAA AAA ATC CGA CGC AAC AAT AGG GCA-3'
<b>K</b>	5'-NH2-C6-AAA AAA AAA AAA ATA ATC TAA TTC TGG TCG CGG-3'
<b>K'</b>	5'-NH2-C6-AAA AAA AAA ACC GCG ACC AGA ATT AGA TTA-3'
<b>L</b>	5'-NH2-C6-AAA AAA AAA AAA AGT GAT TAA GTC TGC TTC GGC-3'
<b>L'</b>	5'-NH2-C6-AAA AAA AAA AGC CGA AGC AGA CTT AAT CAC-3'
<b>M</b>	5'-NH2-C6-AAA AAA AAA AAA AGT CGA GGA TTC GTA ACC TGT-3'
<b>M'</b>	5'-NH2-C6-AAA AAA AAA AAC AGG TTC AGA ATC CTC GAC-3'
<b>N</b>	5'-NH2-C6-AAA AAA AAA AAA AGT CCT CGC TTC GTC TAT GAG-3'
<b>N'</b>	5'-NH2-C6-AAA AAA AAA ACT CAT AGA CGA AGC GAG GAC-3'
<b>O</b>	5'-NH2-C6-AAA AAA AAA AAA ACT TCG TGG CTA GTC TGT GAC-3'
<b>O'</b>	5'-NH2-C6-AAA AAA AAA AGT CAC AGA CTA GCC ACG AAG-3'
<b>P</b>	5'-NH2-C6-AAA AAA AAA AAA ATC GCC GTT GGT CTG TAT GCA-3'
<b>P'</b>	5'-NH2-C6-AAA AAA AAA ATG CAT ACA GAC CAA CGG CGA-3'
<b>Q</b>	5'-NH2-C6-AAA AAA AAA AAA ATA AGC CAG TGT GTC GTG TCT-3'
<b>Q'</b>	5'-NH2-C6-AAA AAA AAA AGA CAC GAC ACA CTG GCT TAT-3'

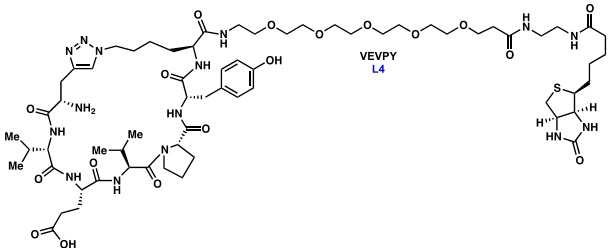
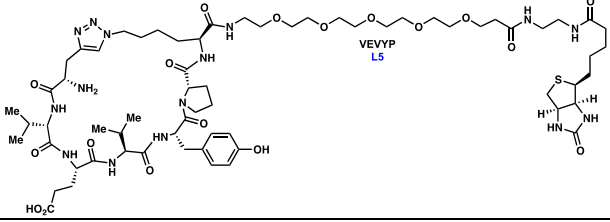
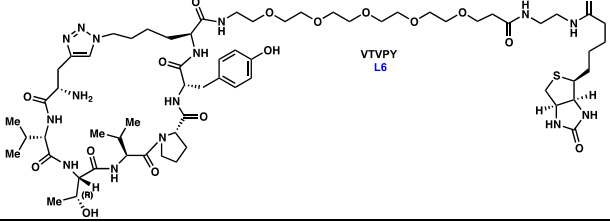
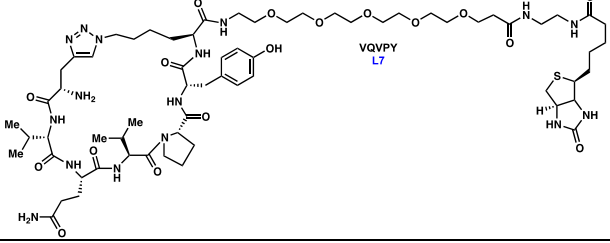
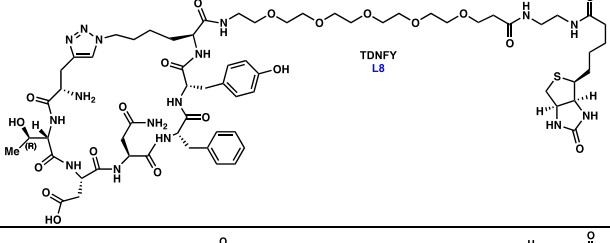
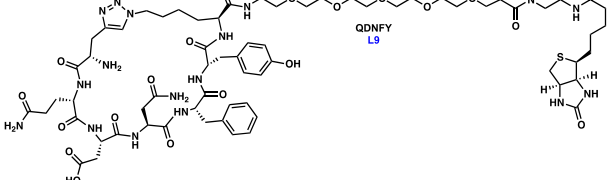
**Table A2C.2. Table of ssDNA sequences used for SAC-DNA conjugation and DNA barcode patterning.** The rows shaded in grey are used for conjugation to SAC protein.

Lane	B'	C'	D'	E'	F'	G'	H'	I'
original scan ( $F_{532}$ )	59024.03	53476.07	54155.37	56759.83	52075.97	51375.13	49365.43	55336.00
six month rescan ( $F_{532}$ )	55831.77	51549.57	50095.27	52715.00	49412.37	48684.57	45213.60	51664.53
Difference ( $F_{532}$ )	3192.27	1926.50	4060.10	4044.83	2663.60	2690.57	4151.83	3671.47
% Difference	5.41	3.60	7.50	7.13	5.11	5.24	8.41	6.63
Lane	K'	L'	M'	N'	O'	P'	Q'	
original scan ( $F_{532}$ )	52048.67	56120.87	52744.73	53185.87	52884.73	52943.00	52287.07	
six month rescan ( $F_{532}$ )	48675.67	53388.87	48104.17	50050.87	49501.07	48818.27	49845.17	
Difference ( $F_{532}$ )	3373.00	2732.00	4640.57	3135.00	3383.67	4124.73	2441.90	
% Difference	6.48	4.87	8.80	5.89	6.40	7.79	4.67	

**Table A2C.3. Analysis of the degradation in barcode quality during extended storage.** The data above is from a barcode validated two months after its preparation and a subsequent re-scan of the validation region six months later. The average difference in  $F_{532}$  output was 3348.80, and the average % difference was 6.26%.

**Table A2C.4. Synthetic epitopes and PCC ligands characterization table.**

Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
	$C_{90}H_{128}N_{20}O_{33}S$	32:30-33:30	[M+H]= 2049.88 [M+Na]= 2071.188	2074.32
	$C_{77}H_{124}N_{20}O_{29}S$	28:30-29:30	[M+H]= 1825.864	1826.22 2
	$C_{90}H_{128}N_{20}O_{33}S$	32:00-33:00	[M+H]= 2049.875 [M+2 Pip-3 OH+H]= 2167.030	2166.39
	$C_{77}H_{124}N_{20}O_{29}S$	25:00-26:00	[M+H]= 1825.649, [M-OH+H] =1809.99 1	1810.51
	$C_{57}H_{95}N_{17}O_{15}S$	3:30-4:30 (f1), 36:00-38:00 (f2)	[M+H]= 1290.699 [M+Na]= 1312.681	1310.81 (f1), 1289.31 (f2)
	$C_{58}H_{97}N_{15}O_{17}S$	2:30-5:00	[M+H]= 1308.699 [M+Na]= 1330.682	1329.9
	$C_{58}H_{97}N_{15}O_{17}S$	3:00-5:00 (f1), 41:30-43:00 (f2)	[M+H]= 1308.699 [M+Na]= 1330.681	1330.16 (f1), 1312.34 (f2)

Structure	Chemical Formula	HPLC Retention Time (min)	Expected mass (amu)	Observed mass (amu)
	$C_{65}H_{103}N_{15}O_{18}S$	36:30 - 39:30 (f1), 40:00 - 41:00 (f2)	[M+H]= 1414.740 [M+Na]= 1436.722	1414.38 (f1), 1435.89 (f2)
	$C_{65}H_{103}N_{15}O_{18}S$	37:00 - 39:00 (f1), 39:00 - 40:00 (f2)	[M+H]= 1414.740 [M+Na]= 1436.722	1413.47 (f1), 1435.88 (f2)
	$C_{64}H_{103}N_{15}O_{17}S$	37:30 - 39:00	[M+H]= 1386.746 [M+Na]= 1408.727	1387.36
	$C_{65}H_{104}N_{16}O_{17}S$	36:00 - 38:30 (f1), 39:00 - 41:30 (f2)	[M+H]= 1413.722 [M+Na]= 1435.738	1412.29 (f1), 1434.80 (f2)
	$C_{66}H_{98}N_{16}O_{20}S$	36:30 - 38:00	[M+H]= 1467.694 [M+Na]= 1489.676	1488.37
	$C_{67}H_{99}N_{17}O_{20}S$	4:30-5:30 (f1), 37:00 - 38:30 (f2)	[M+H]= 1494.705 [M+Na]= 1516.687	1514.14 (f1), 1494.04 (f2)

	Full Lane			Full Centroid			Left Part Centroid			Middle Part Centroid			Right Part Centroid		
Lig.	Avg F <sub>635</sub>	Std-dev	%CV	Avg F <sub>635</sub>	Std-dev	%CV	avg F <sub>635</sub>	Std-dev	%CV	avg F <sub>635</sub>	Std-dev	%CV	avg F <sub>635</sub>	Std-dev	%CV
L6	15968	4436	27.8	17836	3037	17.0	15645	2058	13.2	19382	2578	13.3	18546	3072	16.6
L4a	10434	3093	29.6	11363	2491	21.9	10179	1444	14.2	11181	2208	19.8	12764	2935	23.0
L4b	25079	5502	21.9	27245	4173	15.3	23307	3555	15.3	27398	2128	7.8	31147	2030	6.5
L7	48537	9262	19.1	53496	5310	9.9	48599	4548	9.4	54620	3690	6.8	57411	3113	5.4
L3a	7151	2185	30.6	8117	1291	15.9	7137	1106	15.5	8430	952	11.3	8811	1171	13.3
L3b	15157	3057	20.2	16158	2040	12.6	15070	1916	12.7	16390	2085	12.7	17046	1615	9.5
L5a	7851	2139	27.2	8341	1748	21.0	7918	1268	16.0	8920	1975	22.1	8196	1826	22.3
L5b	31581	5350	16.9	32390	2859	8.8	31947	2750	8.6	32789	2700	8.2	32448	3133	9.7
L7a	32992	7831	23.7	36651	5433	14.8	32020	2774	8.7	37152	4709	12.7	40918	4399	10.8
L7b	14791	4605	31.1	16195	3322	20.5	12956	1819	14.0	16944	1899	11.2	18781	2914	15.5
L9a	16612	4630	27.9	18391	3310	18.0	14966	1107	7.4	20262	2145	10.6	20045	3023	15.1
L9b	41812	7828	18.7	45500	5838	12.8	39380	2637	6.7	47791	5200	10.9	49508	3078	6.2
L1a	18768	3903	20.8	20134	2815	14.0	17838	2497	14.0	20725	2225	10.7	21906	1983	9.1
L1b	22519	4786	21.3	24340	3430	14.1	21889	1763	8.1	24245	2629	10.8	26959	3580	13.3
Bk	14498	3225	22.2	15102	2350	15.6	12609	891	7.1	15486	1525	9.8	17285	1505	8.7

Region	average %CV
Full Lane	23.9
Full centroid	15.5
left part centroid	11.4
middle centroid	11.9
right part centroid	12.3

**Table A2C.5. A pixel by pixel analysis of variance along a barcode lane.** The %CV values for the entire barcode set shown in Figure S15 using the full lane, full centroid, left part of the centroid, middle part of the centroid, and right part of the centroid are displayed in the top table. The average %CV values for each set are displayed in the bottom table.

log[KRAS]	L1a	L1b	L2	L3a	L4a	L4b	L5a
-3.54	28.57	736.24	47.37	38.32	78.34	–	78.36
-3.84	55.65	61.13	73.45	10.80	–	–	23.71
-4.14	63.67	95.71	135.76	61.12	–	129.09	25.89
-4.62	41.32	78.02	37.31	29.92	113.59	18.39	83.24
-5.01	75.01	104.74	91.75	37.41	129.43	126.70	35.44
-5.32	72.72	62.92	311.26	68.47	–	75.72	33.16
-6.01	66.87	58.07	116.58	68.77	129.48	45.88	–
-6.32	43.25	34.28	74.47	117.95	94.72	21.05	114.95
-6.62	27.30	83.53	383.23	219.70	208.89	51.85	93.22
-7.01	91.85	79.90	298.37	112.63	59.01	182.28	88.78
avg %CV	56.62	139.45	156.95	76.51	116.21	81.37	64.08
log[KRAS]	L5b	L6	L7a	L7b	L8	L9a	L9b
-3.54	30.85	20.29	8.43	17.55	2.82	16.76	58.89
-3.84	24.60	80.02	12.77	9.85	3.79	26.82	31.45
-4.14	15.86	73.89	27.32	2.45	2.70	29.67	43.65
-4.62	38.87	–	14.52	10.08	6.82	31.29	34.73
-5.01	41.45	294.01	26.54	28.29	4.31	21.32	34.12
-5.32	29.12	65.36	179.54	59.14	23.74	22.19	38.54
-6.01	52.54	41.04	13.50	87.28	51.17	32.95	42.65
-6.32	69.24	29.65	22.00	178.75	125.33	29.51	70.74
-6.62	152.61	36.53	76.64	79.07	66.46	61.52	36.05
-7.01	16.40	134.31	69.97	16.02	–	30.45	48.13
avg %CV	47.15	86.12	45.12	48.85	31.90	30.25	43.89

**Table A2C.6. The %CV data for the ELISAs.** The %CV measurements are given as the [%CV]. Any %CV>=300% has been marked with an “–”. The global average %CV for the assays was 73.2%.

centroid	L6	L4a	L4b	L7	L3a	L3b	L5a	L5b	L7a	L7b	L9a	L9b	L1a	L1b	Bk
L6	—	***	***	***	***	***	***	***	***	***	n.s.	***	***	***	***
L4a		—	***	***	***	***	***	***	***	***	***	***	***	***	***
L4b			—	***	***	***	***	***	***	***	***	***	***	***	***
L7				—	***	***	***	***	***	***	***	***	***	***	***
L3a					—	***	n.s.	***	***	***	***	***	***	***	***
L3b						—	***	***	***	n.s.	***	***	***	***	**
L5a							—	***	***	***	***	***	***	***	***
L5b								—	***	***	***	***	***	***	***
L7a									—	***	***	***	***	***	***
L7b										—	***	***	***	***	**
L9a											—	***	***	***	***
L9b												—	***	***	***
L1a													—	***	***
L1b														—	***
Bk															—
Full lane	L6	L4a	L4b	L7	L3a	L3b	L5a	L5b	L7a	L7b	L9a	L9b	L1a	L1b	Bk
L6	—	***	***	***	***	*	***	***	***	*	n.s.	***	***	***	**
L4a		—	***	***	***	***	***	***	***	***	***	***	***	***	***
L4b			—	***	***	***	***	***	***	***	***	***	***	***	***
L7				—	***	***	***	***	***	***	***	***	***	***	***
L3a					—	***	**	***	***	***	***	***	***	***	***
L3b						—	***	***	***	n.s.	**	***	***	***	n.s.
L5a							—	***	***	***	***	***	***	***	***
L5b								—	n.s.	***	***	***	***	***	***
L7a									—	***	***	***	***	***	***
L7b										—	***	***	***	***	n.s.
L9a											—	***	***	***	***
L9b												—	***	***	***
L1a													—	***	***
L1b														—	***
Bk															—

Table A2C.7. Calculated p-values for the pixel by pixel full-lane and centroid scatterplots. Symmetric 2-tail, unequal variance p-value matrixes for the centroid and full-lane scatterplots in Figure S15. The p-values are denoted: n.s.  $p > 0.05$ , \*  $0.05 < p < 0.005$ , \*\*  $0.005 < p < 0.0005$ , \*\*\*  $p < 0.0005$ .

	L1a	L1b	L2	L3a	L3b	L4a	L4b	L5a	L5b	L6	L7a	L7b	L8	L9a	L9b	BN-A <sub>20</sub> -Cy3
Fig. 3	0.91	0.92	0.95	0.94	0.81	0.91	0.8	0.97	0.85	0.92	0.95	0.85	0.96	0.94	0.96	N/A
Fig. S16	0.95	0.95	0.96	0.96	0.83	0.93	0.87	0.98	0.88	0.94	0.97	0.89	0.98	0.95	0.97	0.66
Fig. S17	0.56	0.18	0.45	0.73	N/A	N/A	N/A	0.53	0.78	0.73	0.93	0.96	0.97	0.81	0.72	N/A

**Table A2C.8. The goodness of fit measurements for the allosteric KRas binding curves.** The r-squared values for the curves after a double background correction in Figures 3 and S17, and a single background correction in Figure S16.

## Chapter 3

### Development of Epitope-Targeting Protein Catalyzed Capture Agents Against Histidine Rich Protein II

This work described herein is unpublished work that was part of a concerted group effort to develop PCCs against the *unstructured* protein Histidine rich protein II (HRP2), which is a unique biomarker for *Plasmodium falciparum* malaria. These unpublished results (gathered from October 2014 to July 2015) focused on the medicinal chemistry optimization of several hits, and an initial push to develop biligand PCCs. The best ligands (described in JingXin Liang's thesis, **2018**) were characterized for their ability to disrupt the Heme: HRP2 complex *in vitro*, and were found to be potent inhibitors.



## Development of Epitope-Targeting Protein Catalyzed Capture Agents Against Histidine Rich Protein II

### Section 3.1-Introduction

Malarial infection is caused by the *Plasmodium* parasite, and is endemic in the developing regions of the world with warm, humid climates including Sub-Saharan Africa, Asia, and Central/South America. The most virulent form of malaria is called *Plasmodium falciparum* (Pf), and accounts for >90% of all malaria related fatalities.<sup>1</sup> The natural product artemisinin is currently the best drug for treating malarial infections;<sup>2</sup> however, resistance to artemisinin has begun to appear in South Asia. The growing resistance of *Plasmodium* parasites to anti-malarial drug treatments including the “gold standard” drug artemisinin underscores the need to develop new anti-malarial medicines ideally with a unique mechanism of action.

Histidine Rich Protein II (HRP2) is an attractive protein target for treating *Plasmodium falciparum* malarial infection. The parasites utilize the protein to sequester Heme, which is a toxic byproduct of hemoglobin catabolism. HRP2 binds Heme and catalyzes its polymerization into the non-toxic crystalline polymer Hemozoin. Chauhan and coworkers determined that HRP2 has two binding modes for Heme that are pH-dependent.<sup>3</sup> Below pH=6, when the histidine (His) side chain is in its acidic form, the carboxylate side chains of aspartate (Asp) and glutamate (Glu) bind Heme, but, above pH=6, binding to the histidine side chain dominates. Chauhan also demonstrated that Heme polymerization is primarily catalyzed by the carboxylate-binding mode, which is consistent with HRP2 sequestration in the acidic food vacuoles of the parasite.<sup>3</sup> Artemisinin's mechanism of action inhibits Hemozoin formation by first covalently

binding to Heme before HRP2 complexes with the Art-Heme moiety.<sup>4-6</sup> The Art-Heme complex cannot undergo polymerization, and consequently Heme builds up and eventually kills the *Plasmodium* parasite. An alternative mechanism to inhibit Hemozoin formation would involve binding to the HRP2 protein and preventing it from complexing with Heme molecules. As this is a novel inhibition pathway this could lead to new therapeutics that treat Pf malarial infection.

While an attractive target, HRP2 presents unusual challenges for developing molecules to bind HRP2 protein. First, the HRP2 sequence primarily consists of repeating sequences of three amino acids: alanine (A), aspartic acid (D), and histidine (H). This results in a very “sticky” polypeptide chain that lacks a defined 2°/3° structure. Consequently, HRP2 does not have any well-defined grooves or pockets that a small molecule could bind. Secondly, fourteen unique epitopes have been identified in HRP2 samples that appear in varying quantities across different

**Table 3.1. The fourteen unique epitopes found in HRP2.**

Code	Repeat Sequence
1	AHHAHHVAD
2	AHHAHHAAD
3	AHHAHHAAY
4	AHA
5	AHHAAHHASD
6	AHHATD
7	AHHAAD
8	AYY
9	AHHAAAHHATD
10	AHHAAAHHATD
11	AHN
12	AHAAAHHHEAATH
13	AHHASD
14	AHHAHHATD

geographical regions (Table 3.1).<sup>7</sup> The extensive variation in epitope prevalence in Table 3.2 highlights the unusual amount of sequence diversity in HRP2. This makes developing reagents to bind HRP2 in an epitope-targeted manner a challenge as targeting a single epitope could result in ineffective binding in geographic regions where that epitope is either absent or in low abundance.<sup>8</sup>

**Table 3.2. The prevalence of the different HRP2 epitopes in patient samples from endemic areas. Taken from reference 8.**

Region, country/island of origin	No.	Total	No. of total and individual repeats													
			1	2	3	4	5	6	7	8	9	10	11	12	13	14
Africa																
Burkina Faso	1	30	1	12	0	1	0	3	8	2	0	2	0	1	0	0
Cameroon	2	32-35	5	11-12	1-2	0	1	4-5	4-8	1	0	1-2	1-2	1	0-1	0-1
Gambia	1	34	2	9	1	0	2	7	9	2	0	1	0	1	0	0
Ghana	2	31-33	1-3	12-14	1-3	0-1	0-1	3-4	6-8	1	0	1-2	0	1	0	0
Liberia	1	33	3	12	1	2	1	3	8	1	0	1	0	1	0	0
Nigeria	1	33	1	9	1	0	2	7	9	2	0	1	0	1	0	0
Sudan	3	28-31	1-5	7-11	1	1	1-2	2-6	5-9	1-2	0	0-2	0	1	0	0
Uganda	1	29	3	13	2	0	0	2	7	1	0	0	0	1	0	0
Zambia	1	31	3	13	2	0	1	3	6	1	0	1	0	1	0	0
Pacific																
Indonesia	2	33	1	9	1	0	2	7	9	2	0	1	0	1	0	0
Malaysia	1	29	3	13	2	0	0	2	7	1	0	0	0	1	0	0
Philippines	21	25-35	1-4	6-15	0-3	0-3	0-2	2-4	1-10	1-2	0-1	0-2	0-1	1	0	0
Papua New Guinea	9	24-38	1-3	7-15	1-2	0-2	1-3	2-7	2-9	1-2	0	1-2	0	1	0	0-1
South America																
Brazil	5	31-33	2-4	9-12	2	0	0-2	2-5	8-10	0-1	0	0-3	0	1	0	0
Ecuador	1	33	5	12	1	0	1	4	7	1	0	1	0	1	0	0
Honduras	1	32	5	12	1	0	1	2	8	1	0	1	0	1	0	0
Peru	1	32	4	13	1	1	0	2	6	1	0	2	0	1	1	0
Southeast Asia																
Cambodia	1	29	3	13	2	0	0	2	7	1	0	0	0	1	0	0
Vietnam	5	30-36	2-6	13-15	1-2	1-2	1	3-4	2-7	1	0	1-2	0	1	0	0
Solomon Islands	7	30-36	2-4	11-16	1-2	0-1	1-2	3-7	4-9	1	0	1-2	0-1	1	0	0-1
Thailand	7	27-36	1-5	10-19	1-3	0-3	0-1	2-4	1-6	0-2	0	0-2	0	1	0-1	0-1
Overall		25-38	1-6	6-19	0-3	0-3	0-3	2-7	1-10	0-2	0-1	0-3	0-2	1	0-1	0-1

However, PCC agents can be easily raised against a variety of epitopes in a target protein then linked together to form a biligand that can target multiple epitopes. This would mitigate issues with detecting HRP2 as a result of its sequence diversity; also, a multiligand would enhance the binding affinity through cooperativity to enable detection/inhibition of HRP2 at low physiological levels of HRP2 protein. The remainder of this chapter summarizes efforts to develop PCC agents against the HRP2 protein.

## Section 3.2-Methods<sup>9</sup>

### 3.2.1-Preparation of Peptides<sup>10</sup>

The peptides were prepared using standard Fmoc/Bu coupling procedures either by hand or using an automatic synthesizer. Macrocyclization occurred via a copper-catalyzed click reaction before simultaneous resin cleavage and global deprotection of side chains using a 95% trifluoroacetic acid (TFA) cleavage cocktail. Precipitation of the

crude peptide into diethyl ether preceded purification using semi-preparative high-performance liquid chromatography (semi-prep HPLC). Fractions containing pure peptide were identified using matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry and lyophilized (For characterization spectra see **Figures A3B.1-A3B.37**). The resulting powder was reconstituted in dimethyl sulfoxide (DMSO), quantified using Nanodrop, and stored at 4 °C until use.

### ***3.2.2-Enzyme-Linked Immunosorbent Assay Analysis of Peptides***

The overnight ELISA protocol commenced with immobilizing the biotinylated peptides and blank on a Neutravidin coated 96-well plate before blocking the plate (with 5% milk or 3% BSA) overnight at 4 °C. Protein incubation with a GST-tagged recombinant HRP2 protein also occurred for 7-8 hours at 4 °C before incubation with an anti-GST HRP-conjugated antibody (ab) at RT. The assay was developed for several minutes at RT using a TMB peroxidase development kit before quenching with 1 M H<sub>2</sub>SO<sub>4(aq)</sub> and immediately measuring the absorbance using a Flexstation 3 plate reader. The data was double background corrected before graphing using Graphpad Prism 7.

The all-room temperature ELISA protocol mirrored the overnight ELISA protocol except that the blocking and protein incubation steps occurred at RT for two hours each. The data was double background corrected as before and graphed using Graphpad Prism 7.

### ***3.2.3-Linker Library Synthesis***

Linker libraries were developed using either the anchor **JXL1** or **DNB**. Linker synthesis occurred by an automated five cycle split and mix protocol using the amino acids pool Fmoc-D-Leu-OH, Fmoc-Gly-OH, Fmoc-D-Pro-OH, Fmoc-protected  $\alpha$ -amino isobutyric acid (Fmoc-Aib-OH), or nothing. The resulting library had a variable linker

with between zero and five amino acids. A Boc-protected propargyl click handle was coupled on before global deprotection of all acidic protecting groups. Each library was stored in 0.05% (w/v) sodium azide in MQ H<sub>2</sub>O at 4 °C until use.

#### ***3.2.4-Linker Library Screen Pre-Clear***

For the pre-clear step each library was split into three aliquots and screened individually against the following biotinylated ligands: JXL1 library **DNB**, **N-term**, and **C-term** ligands; DNB library **JXL1**, **N-term**, and **C-term** ligands. After swelling the library in 0.1% BSA binding buffer and blocking overnight with 1% BSA blocking buffer the library was incubated with one of the anchors for one hour at RT. Incubation with a mouse anti-biotin alkaline phosphatase-conjugated antibody preceded stripping with a high salt buffer over one hour and development with the BCIP/NBT development system for thirty minutes. The development reaction was quenched with 6 M hydrochloric acid and all dark purple beads were removed before decolorizing the remaining beads with NMP and stripping with 7.5 M guanidinium chloride (Guan•HCl), pH = 2.0.

#### ***3.2.5-Linker Library Product Screens***

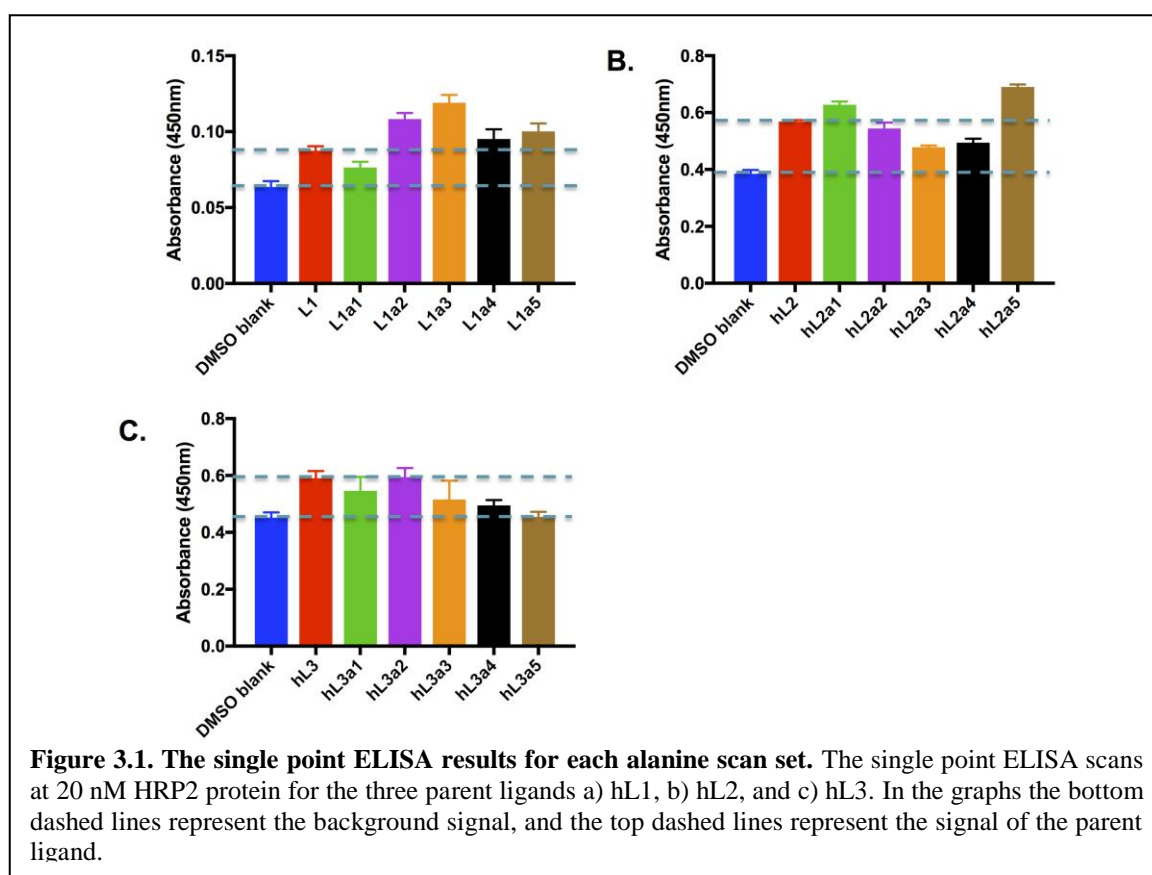
Each aliquot was swelled in binding buffer and blocked overnight with 1% BSA before incubating with recombinant HRP2 protein and the biotinylated ligand used in the pre-clear stage. Any non-covalently bound material was stripped from the beads using 7.5 M Guan•HCl, pH=2.0 for one hour and incubation with a high salt buffer for an additional hour. Development occurred in the same manner as the pre-clear, and the dark purple beads were picked out and sequenced using Edman degradation.

### **Section 3.3-Results and Discussion**

#### ***3.3.1-Alanine Scan of Sequence Hits***

Previous work performed in the Heath group included combinatorial click library

screens of several epitopes that are common in HRP2 samples: the shifted Type 2 epitope AHHAADAHH, the N-term epitope LHETDQAHVDD, the C-term epitope AHHATDAHHAAAHHEAATH, and the Type 6 epitope AHHATD. These screens yielded several dozen first generation hits. Each hit needed to be tested for binding to the full protein before the best binders were optimized to maximize their binding affinity. My research focused on optimizing the structures and identifying a best binder with three hits from the shifted type 2 epitope screen: YYQLL (**hL1**), YYYNV (**hL2**), and LGYYY (**hL3**).



The initial round of medicinal chemistry optimization involved mapping out the binding contributions of the residues of each peptide with alanine substituted variants. Any alanine variants with dramatically different binding to HRP2 protein would have that position selected for further amino acid substitutions. Each alanine scan set was prepared

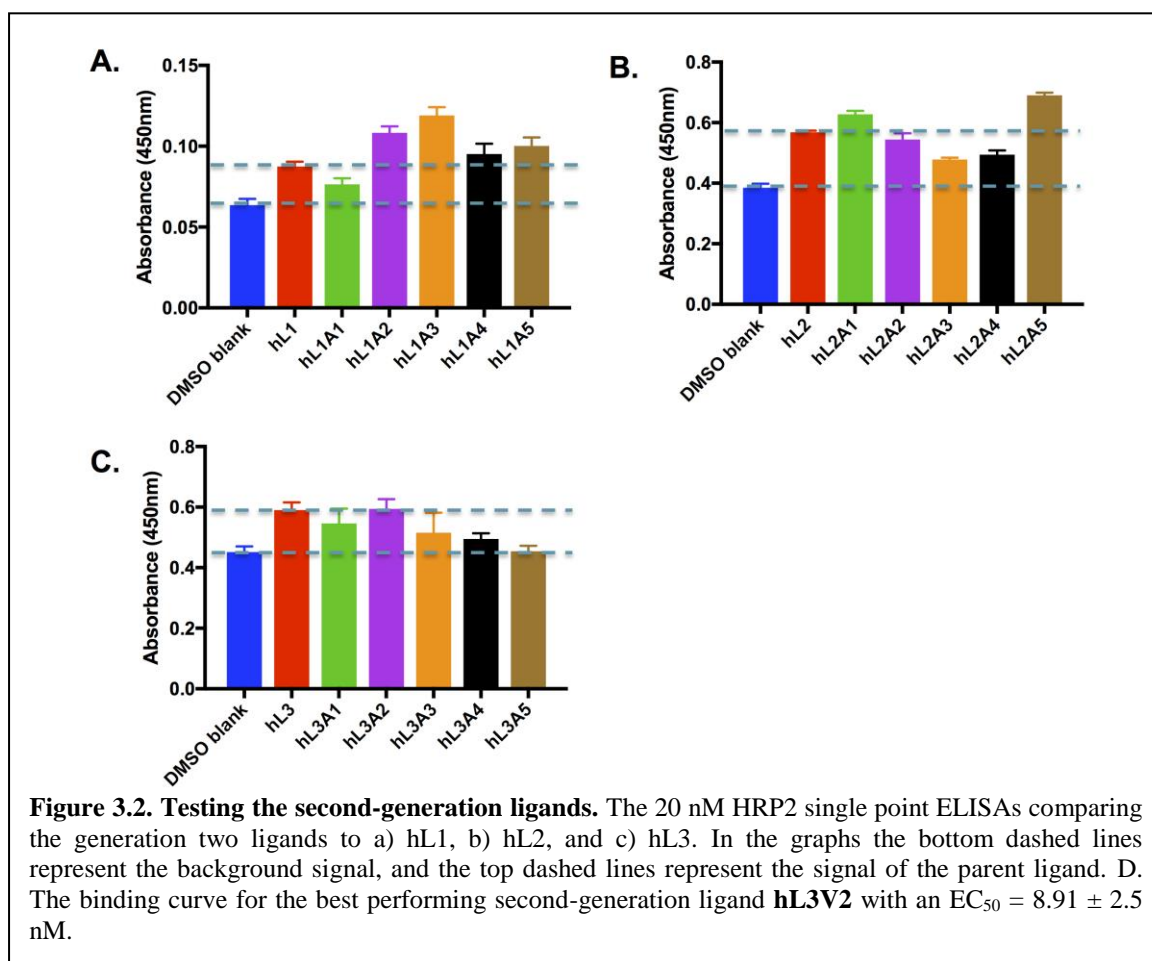
(Table A3C.1) and evaluated against the parent ligand in a single point sandwich ELISA (Figure 3.1). The scan for ligand **hL3** revealed that the tyrosine residues in positions 3-5 strongly bind HRP2. Swapping in structurally similar ligands at “intolerant” positions presents a unique opportunity to probe the non-covalent interactions involved in HRP2 binding while potentially improving affinity. Thus, the strongly intolerant residues of **hL3** were also chosen for substitution.

### *3.3.2-Second-Generation Ligand Synthesis and Measurement of HRP2 Binding Affinity*

The individual alanine scans yielded between one and four residues for substitution respectively, and the residues to be modified were Glycine (G), Tyrosine (Y), and Glutamine (Q). Each residue's characteristics were analyzed to identify appropriate substitutions. For example, Gly lacks a side chain, and consequently does not impart any steric crowding to a peptide sequence; also, Gly makes the macrocyclic peptide ring more conformationally flexible at that position. An increase in HRP2 binding with the addition of a small, but sterically important methyl side chain in the Ala scan suggests that a more rigid ring could lead to an increase in binding affinity. Thus, Val was selected as an amino acid residue substitution for Gly. Tyrosine has an aromatic phenol chain whose hydroxyl group can accept and donate hydrogen bonds. Also, the hydroxy group uses resonance to increase the  $e^-$  density of the aromatic ring overall. Exchanging the  $-OH$  group for a fluorine atom could result in stronger H-bonds with the His residues in HRP2. Also, the small size and increased electronegativity of fluorine results in an electron-deficient aromatic ring. A reduction in electron density could increase the strength of anion- $\pi$  interactions with the aspartate residues on HRP2. This effect could be enhanced with additional F on the same aromatic ring so both 4-fluorophenylalanine ( $^4F$ ) and 3,4-difluorophenylalanine ( $^{3,4}F$ ) were chosen to replace Tyr. Glutamine is the uncharged

amide analogue of the charged amino acid glutamate (Glu). Exchanging an uncharged residue for an analogous charged residue could enhance the binding affinity with HRP2, due to more salt bridge interactions with HRP2. Thus, both Asp and Glu were chosen to replace Gln.

The second-generation ligand sets were prepared (Table A3C.2) and evaluated against the parent ligand as before (Figure 3.2). The Generation II **hL1** single-point ELISA



revealed that exchanging the uncharged amide residue for a charged acidic residue did not result in any change in the binding affinity (Figure 3.2A). The His residues ( $pK_a = 6.0$ ) are uncharged in the binding buffer ( $pH = 7.4$ ), so adding a residue that would encourage salt bridge interactions with the acidic form of His was not productive.



Glutamine could also have been substituted with arginine (Arg), but since **hL1** was the weakest binder of the three ligands studied, further modifications were not pursued. The Generation II **hL2** ELISA (Figure 3.2B) did result in enhanced signal with the peptide **hL2<sup>f</sup>F2**, but as it was not as good a binder as the Y<sup>4F</sup>FYRV (**JXL1** ligand) sequence developed by a coworker, this sequence was not subjected to further optimization. The ELISA for the second-generation **hL3** ligands revealed that the addition of fluorinated analogues at the Tyr residues did not result in a significant increase of HRP2 binding (Figure 3.2C). This suggests that the cation- $\pi$  interactions between the weakly charged His residues and the Tyr side chains play a strong role in **hL3** binding HRP2. Tyr residues appearing in most of the original hits from the shifted Type 2 epitope click library screen further highlighted tyrosine's significance. Surprisingly, the substitution of Gly for Val in ligand **hL3V2** resulted in a dramatic increase in signal, which suggests that ring rigidity played a larger role in HRP2 binding affinity than previously expected. A full binding curve for **hL3V2** was prepared, and it was found to have an EC<sub>50</sub> of 8.91 $\pm$ 2.5 nM (Figure 3.2D). The ligand **hL3V2** was the only ligand carried onwards for further optimization.

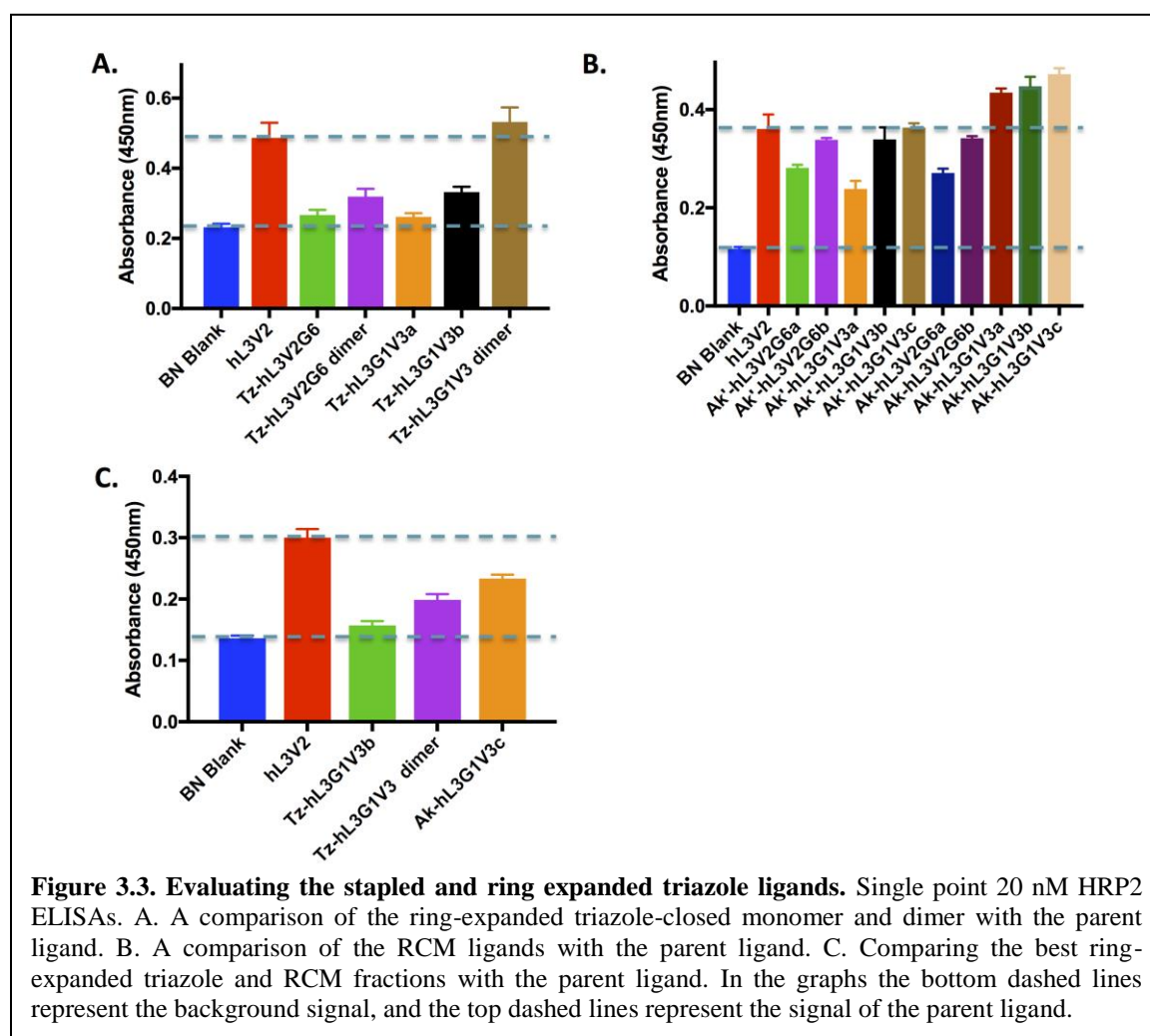
### 3.3.3-Further Optimization of Ligand Hit **hL3V2**

The next round of optimization involved investigating the ideal ring size and closure for the macrocycle. The ligands to date have all been 27-atom macrocycles. A slightly larger macrocyclic ring could result in a binding pocket that better fits the shifted Type 2 HRP2 epitope. A ring-closing metathesis (RCM) ring closure, which gives a ring size of 33 atoms, is an unnatural ring closure that utilizes the non-canonical stapled peptides Fmoc-(S)- $\alpha$ -Me-2'-(7-octenyl)-alanine (Fmoc-Oct-OH) and Fmoc-(R)- $\alpha$ -Me-2'-(4-pentenyl)-alanine (Fmoc-Pen-OH).<sup>11,12</sup> A pentameric peptide sequence that is ring closed

with stapled peptides does require adding a Gly residue to one end of the sequence in order to satisfy the  $i, i + 7$  or  $i, i + 4$  rule.<sup>13</sup> As the RCM closure represents a change in ring flexibility, due to the introduction of quaternary centers and an increase in the hydrophobicity with the stapled peptide side chains, the RCM closure was chosen for ring size/closure optimization. A larger binding pocket could also spontaneously result from the click reaction used to close the macrocycle. The macrocyclic ring closure is the major product of the click reaction, but a side product can result from an *intermolecular* click reaction between two peptide sequences. The resulting ligand would possess two triazole closures for a ring size of 54 atoms (Figure A3B.38). The Type 2 epitope binder YKYR (**DNB** ligand), which was identified by a colleague, had increased binding affinity for HRP2 in its dimer form.<sup>10</sup> It has been shown in the literature that longer sequences favor intermolecular click reactions as a result of an increase in conformational flexibility.<sup>14</sup> Adding a Gly residue to encourage dimerization also increases the triazole closed macrocyclic ring size up to 30 atoms, which would allow for testing if adding a conformationally flexible residue encouraged dimerization without depressing HRP2 binding affinity; also, the added Gly would provide a triazole analogue to compare to the RCM closed ligands. This comparison would determine whether the change in binding affinity was due to the larger ring size or to the change in macrocyclic ring closing. As a result the following isomers of the ring expanded and RCM-closed ligand variants of **hL3V2** were prepared (Table A3C.3).

Although both the heterodimer and homodimer are possible side products of the RCM ring closing, only monomer was observed by MALDI-TOF analysis. Each ligand was isolated in multiple fractions, which results from the Grubbs I catalyst producing a

mixture of E/Z alkene isomers. The different ligand fractions were tested, and the best isomer was **Ak-hL3G1V3** (Figure 3.3). This particular isomer has the stapled peptide with the shorter linker at the C-terminus and places the extra Gly residue away from the Tyr residues. The Tyr residues are heavily involved in binding HRP2, which suggests that optimal binding requires conformational rigidity near the residues that actively bind HRP2.



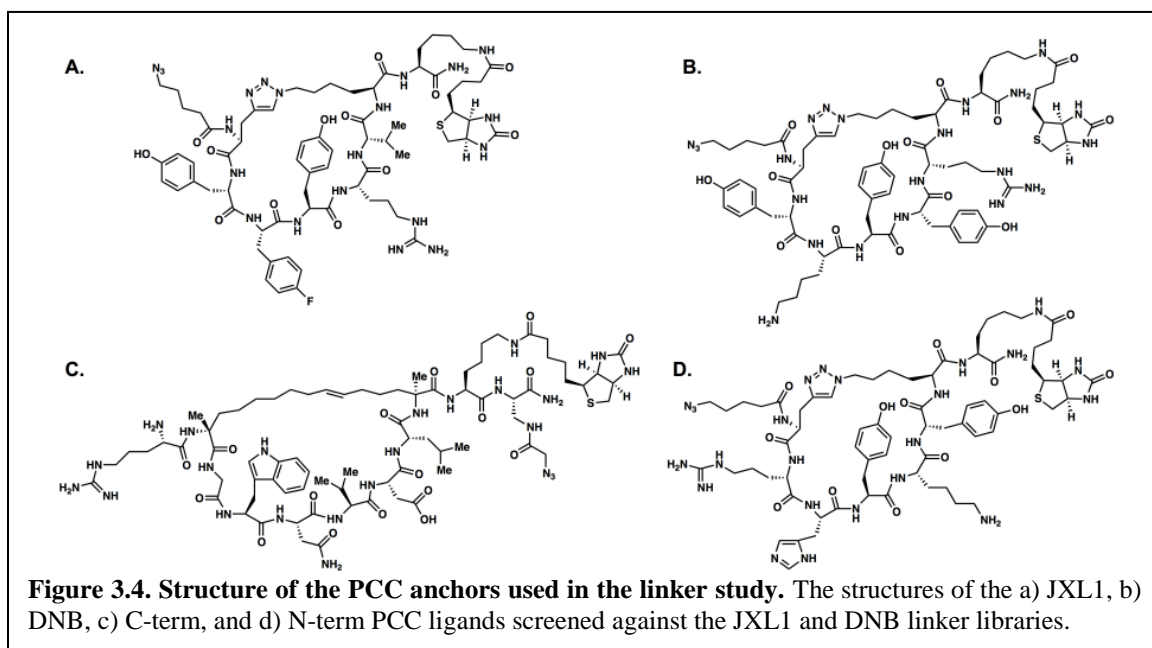
Preparing the ring-expanded triazole ligands did lead to the isolation of a significant amount of dimer in addition to the monomeric ligands. As expected, the dimeric ligands outperformed the monomeric ligands, likely as a result of cooperativity (Figure 4.3B).

However, dimerizing the ring-expanded ligands did not result in a marked improvement compared to the parent ligand **hL3V2**. The drop in binding affinity, despite having two copies of the HRP2 binding sequence present, could result from the extra Gly residues. Although having the Gly away from the active Tyr residues is preferred, the extra Gly may introduce enough conformational flexibility in the dimer's ring to disfavor HRP2 binding. This suggests that small, conformationally rigid cycles are needed to bind HRP2 selectively. Directly comparing the best RCM ligand to the best ring-expanded triazole did show an increase in binding affinity for the RCM peptide, but overall both ligands behaved worse than the parent ligand **hL3V2** (Figure 3.3C). This reinforced that, for targeting the shifted type 2 epitope, the original macrocycle ring size and ring closure was optimal.

### ***3.3.4-Developing PCC Agent Biligands***

Although the capture agents developed to bind select HRP2 epitopes have EC<sub>50</sub> values in the 10-20 nM range, the best antibodies against HRP2 have much lower K<sub>d</sub> values: 1 nM (C1-13 epitope, IgG antibody) and 160 pM (PTL3 epitope, IgM antibody).<sup>15</sup> The difference in binding affinity between the antibodies and the much smaller capture agents can be explained by cooperativity. Improving the EC<sub>50</sub> of the PCC agents further would require linking multiple copies of the cyclic peptides together. One such method would be to prepare a biligand composed of two monomeric PCC agents coupled through a spacer that can either be a flexible PEG<sub>n</sub> molecule or a rigid polypeptide linker. Thus, work was undergone collaboratively to identify the optimal amino acid spacer for linking together the best Type 2, shifted Type 2, N-term, and the C-terminal epitope targeting monoligands. Two linker libraries were prepared with the ligands **JXL1** (EC<sub>50</sub> = 538 pM) and **DNB** (type 2 epitope, EC<sub>50</sub> = 218±44 nM) as anchors that were coupled to linkers

that were 0-5 amino acids in length. The amino acid pool for linker library split and mix synthesis was: glycine (G), D-proline (p), D-leucine (l), 2-amino isobutyric acid (Aib), or nothing. These linker libraries were individually screened against the following biotinylated ligands: **DNB**, **JXL1**, **N-term** ( $EC_{50} = 4$  nM), and **C-term** (Figure 3.4).



The linker screens generated a total of 25 hits (Table A3C.4). Two of the hits (**DNBvCtermH6** and **DNBvCtermH8**) were evaluated in a double-point ELISA against the monoligands and were found to outperform the anchors. The **DNBCtermH8** biligand was found to have an  $EC_{50} = 540$  pM, which represents a forty-fold increase over the  $EC_{50}$  value of the original DNB anchor.

At this stage the project leader wrapped up the HRP2 protein study by examining ligands' ability to prevent complexation between free Heme and HRP2 protein. The functional study revealed that the **N-term**, **DNB**, and **JXL1** ligands all inhibited the complexation, but that the **JXL1-PEG-DNB** biligand inhibited Heme: HRP2 complex formation even better than chloroquine, which is a commercially available malaria

prophylactic. The results of this functional study are summarized both in a recent publication under review<sup>16</sup> and in JingXin Liang's dissertation.

### Section 3.4-Conclusion

In closing, this report disclosed efforts to optimize the library screen hits **hL1**, **hL2**, and **hL3** to arrive at the ligand **hL3V2**. Further studies on changing macrocyclic ring closure and macrocyclic ring size revealed that low-conformational flexibility and a smaller ring size are optimal for binding the shifted Type 2 epitope AHHAADAHH. The best ligands discovered during the group effort to test and optimize all of the HRP2 epitope screen hits were tested for inhibition of the Heme: HRP2 complex and found to inhibit at a rate comparable to or better than chloroquine, a common anti-malarial.

### Section 3.5-References

- (1) WHO. WHO | Malaria <http://www.who.int/mediacentre/factsheets/fs094/en/> (accessed Feb 21, 2018).
- (2) Malaria Consortium. Malaria Consortium - Artemisinin-based Combination Therapy (ACT) (Pages) <http://www.malariaconsortium.org/pages/112.htm> (accessed Feb 21, 2018).
- (3) Lynn, A.; Chandra, S.; Malhotra, P.; Chauhan, V. S. *FEBS Lett.* **1999**, 459 (2), 267.
- (4) Accardo, A.; Laurent, S. A.-L.; Mazarguil, H.; Meyer, M.; Robert, A.; Meunier, B. *J. Inorg. Biochem.* **2007**, 101, 1739.
- (5) Pandey, A. V.; Babbarwal, V. K.; Okoyeh, J. N.; Joshi, R. M.; Puri, S. K.; Singh, R. L.; Chauhan, V. S. *Biochem. Biophys. Res. Commun.* **2003**, 308, 736.
- (6) Kannan, R.; Sahal, D.; Chauhan, V. S. *Chem. Biol.* **2002**, 9, 321.
- (7) Baker, J.; McCarthy, J.; Gatton, M.; Kyle, D. E.; Belizario, V.; Luchavez, J.; Bell,

- D.; Cheng, Q. *J. Infect. Dis.* **2005**, *192*, 870.
- (8) Lee, N.; Baker, J.; Andrews, K. T.; Gatton, M. L.; Bell, D.; Cheng, Q.; McCarthy, J. *J. Clin. Microbiol.* **2006**, *44*, 2773.
- (9) *For detailed protocols of all experiments see the supporting information file.*
- (10) Das, S.; Nag, A.; Liang, J.; Bunck, D. N.; Umeda, A.; Farrow, B.; Coppock, M. B.; Sarkes, D. A.; Finch, A. S.; Agnew, H. D.; Pitram, S.; Lai, B.; Yu, M. B.; Museth, A. K.; Deyle, K. M.; Lepe, B.; Rodriguez-Rivera, F. P.; McCarthy, A.; Alvarez-Villalonga, B.; Chen, A.; Heath, J.; Stratis-Cullum, D. N.; Heath, J. R. *Angew. Chemie Int. Ed.* **2015**, *54*, 13219.
- (11) Schafmeister, C. E.; Po, J.; Verdine, G. L. *J. Am. Chem. Soc.* **2000**, 5891.
- (12) Brik, A. *Adv. Synth. Catal.* **2008**, 1161.
- (13) Blackwell, H. E.; Sadowsky, J. D.; Howard, R. J.; Sampson, J. N.; Chao, J. A.; Steinmetz, W. E.; O'Leary, D. J.; Grubbs, R. H. *J. Org. Chem.* **2001**, *66*, 5291.
- (14) Turner, R. A.; Oliver, A. G.; Lokey, R. S. *Transform* **2007**, No. 13, 1.
- (15) Kattenberg, J. H.; Versteeg, I.; Migchelsen, S. J.; González, I. J.; Perkins, M. D.; Mens, P. F.; Schallig, H. D. F. H. *MAbs* **2012**, *4*, 120.

## **Appendix 3A**

### **Materials and Experimental Procedures**

#### **Table of Contents**

<b>Materials</b>	115
<b>Methods</b>	
General Procedures for the Preparation of Ligands	118
Overnight ELISA Procedure	121
Room Temperature ELISA Procedure	122
Linker Library Pre-Clear	123
Linker Library Product Screen	125
<b>References</b>	125



## Materials

Unless otherwise stated all chemicals were used as received.

### Peptide Synthesis

The 9-Fluorenylmethoxycarbonyl (Fmoc) and t-butyl carboxy (Boc) protected amino acids were purchased from Anaspec, AAPTEC, Bachem, ChemPep, and Sigma-Aldrich. The Fmoc-protected stapled amino acids Fmoc-(S)-2-(7-octenyl) Ala-OH and Fmoc-(R)-2-(pentenyl) Ala-OH used to prepare the RCM-closed peptides were purchased from Sigma Aldrich. Biotin NovaTag™ resin was obtained from EMD Chemicals, Inc. and used for synthesis of biotinylated peptides. The Seiber amide resin used for the synthesis of protected peptide fragments and the Rink Amide MHBA Resin used for the preparation of C-term amide peptides were purchased from AAPTEC. The peptide one-bead-one compound (OBOC) library was prepared on Tentagel Resin purchased from RAPP Polymere. The Fmoc-protected propionic acid polyethylene glycol (PEG<sub>n</sub>) linkers were purchased from ChemPep Inc. The L-Ascorbic Acid and copper (I) iodide (CuI) used for click reactions were purchased from Sigma Aldrich. The N-methyl pyrrolidine (NMP), 1-[Bis (dimethylamino) methylene]-1H-1, 2, 3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate (HATU), N, N'-dimethyl formamide (DMF), and N, N'-diisopropylethylamine (DIPEA) used in peptide synthesis were bought from EMD Chemicals, Inc., ChemPep, and Sigma-Aldrich respectively. The piperidine used for the removal of Fmoc protecting groups was purchased from Sigma-Aldrich. The Hydrazine (60 (% (v/v) in H<sub>2</sub>O) and 2-azido acetic acid used for the deprotection of 1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)-3-ethyl (Dde) protecting groups and formation of Az4' groups were purchased from Sigma Aldrich. The Benzyldiene-bis

(tricyclohexylphosphine)-dichlororuthenium (Grubbs I) catalyst and 1,2-dichloroethane (DCE) used for the preparation of stapled peptides were purchased from Fischer Scientific.

### **Peptide Isolation and Purification**

The trifluoroacetic acid (TFA), and triethylsilane (TESH) used for the cleavage of peptides from resin and global deprotection of side chains were purchased from Sigma-Aldrich. The diethyl ether used to precipitate crude peptide was purchased from JT Baker. The Omnisolv grade acetonitrile (MeCN) used for peptide purification was purchased from EMD Millipore. Unless otherwise stated, peptide preparation was performed using a Titan 357 Automatic Peptide Synthesizer (AAPPTec, Louisville, KY). Mass analysis was performed using a Voyager De-Pro matrix assisted laser desorption ionization time-of-flight mass spectrometer (MALDI-TOF MS) (Applied Biosystems, California). The crude peptides were dissolved in either DMSO (Sigma Aldrich) or (1:1) MeCN/doubly distilled water (MQ H<sub>2</sub>O) w/ 0.1% TFA before purification by a gradient of 0% to 50% acetonitrile in MQ H<sub>2</sub>O with 0.01% (v/v) TFA using a RP-HPLC (Beckman Coulter System Gold 126 Solvent Module and 168 Detector) using a C18 reversed phase semi-preparative column (Phenomenex Luna 10 µm, 250 × 10 mm). The concentration of peptides was determined using a Nanodrop 2000 Spectrophotometer (ThermoFischer Scientific Inc., Massachusetts).

### **ELISA Analysis of Peptides**

The bovine serum albumin (BSA) and nonfat milk powder used in ELISAs was purchased from Sigma-Aldrich and Best Value respectively. The Glutathione S-transferase (GST)-labeled recombinant *P. falciparum* antigen HRP2 protein that was used

in the ELISAs was purchased from CTK Biotech Inc. The anti-GST HRP-conjugated antibody (ab58626) was purchased from Abcam. The ELISAs were run on either 96-well clear Pierce Neutravidin Plates (#15129) or Pierce Neutravidin Coated Plates (#15127) purchased from ThermoFischer Scientific. The tris (hydroxymethyl) aminomethane hydrochloride salt (Tris•HCl), sodium chloride, and magnesium chloride pentahydrate ( $\text{MgCl}_2 \cdot 5\text{H}_2\text{O}$ ) used to prepare the Tris buffer were purchased from Sigma Aldrich. The surfactant polysorbate 20 (Tween20) was purchased from Sigma Aldrich. The TMB Microwell Peroxidase Substrate System B (#50-76-00) that was used to develop ELISAs was purchased from KPL. The sulfuric acid ( $\text{H}_2\text{SO}_{4(\text{aq})}$ ) used to quench the enzymatic amplification reaction in the ELISAs was purchased from JT Baker. The 96-well ELISA plates were read using a Flexstation 3 plate reader (Molecular Devices LLC, Sunnyvale, CA).

### **Linker Library Screens**

The recombinant, unlabeled HRP2 protein used for the linker screens was a generous gift from the Program for Appropriate Technology in Health (PATH). The mouse anti-biotin-alkaline phosphatase conjugated antibody (ab) (#A6561) pre-clear and the subsequent product screens were purchased from Sigma Aldrich. The combined 5-bromo-4-chloro-3-indoyl phosphate (BCIP)/ nitro blue tetrazolium (NBT) (#S3771) used to develop hits during the library screens was purchased from Promega. The concentrated hydrochloric acid used to quench the BCIP/NBT development was purchased from Sigma Aldrich. The methanol (MeOH) used to rinse the library beads was purchased from EMD Millipore. The guanidinium chloride (Guan•HCl) used to strip library beads was purchased from Sigma Aldrich. The sodium azide preservative was purchased from

Sigma Aldrich. Sequencing of bead hits occurred via Edman degradation sequencing on a Procise Protein Sequencer (Applied Biosystems, California).

## **Methods**

### ***General Procedures for the Preparation of Ligands***

#### **Preparation of Resin for Coupling**

Novatag Biotin Resin (0.44 mmol amine/g, 300 mg, 0.132 mmol) was placed into a 8mL solid-phase peptide synthesis (SPPS) tube fitted with a stopcock and swelled in 4mL of NMP for a minimum of two hours.

#### **Hand Deprotection of Fmoc Protecting Group**

The resin was incubated with 4mL portions of Piperidine (20% (v/v) in NMP) for 2 min, 5 min, and 25 min before rinsing with NMP (5x).

#### **Peptide Hand Couplings**

The appropriate Fmoc-protected amino acid (2.00-4.00 equiv.) was added directly to the SPPS tube with HATU (2.00-4.00 equiv.). After solubilizing the coupling reagents in NMP, the solution was charged with DIPEA (6.00-12.0 equiv.) and allowed to react on a peptide shaker at room temperature (35 min → overnight) before washing with NMP (5x).

#### **Titan Synthesis**

Automated synthesis of the 5-mer and alkyne click handle was conducted on the Titan 357 Automatic Peptide Synthesizer. Wells were charged with 50mg each of Novatag resin following Fmoc-Az4-OH coupling. Titan synthesis prepared the peptides in a linear fashion: Fmoc-NH-Pra-X<sub>1</sub>X<sub>2</sub>X<sub>3</sub>X<sub>4</sub>X<sub>5</sub>-Az4-PEG<sub>5</sub>-Resin.

#### **Click Reaction**

The click reaction solution for each ligand (50mg, 0.0220 mmol) was prepared as a solution in 3mL of Piperidine (20% (v/v) in NMP): CuI (8.4 mg, 0.0440 mmol) and L-ascorbic acid (19.4 mg, 0.110 mmol). The solution was added to the resin and allowed to cyclize overnight with shaking at room temperature. A solution of click wash was prepared immediately before use: DIPEA (5% (v/v)) and sodium diethyl dithiocarbonate (5% (w/v)) in NMP. After draining the SPPS tube the resin was rinsed with click wash (3x), and incubated with click wash for ten minutes before rinsing with NMP (5x).

### **Ring-Closing Metathesis of Macrocyclic Peptide Ring**

The resin (200 mg, 0.0880 mmol) was rinsed with DCM (5x) and left to dry for a minimum of one hour before reswelling in 1,2-DCE (4 mL) for one hour. An oven-dried 2-neck round-bottom flask (RBF) under Ar<sub>(g)</sub> was charged with the resin solution and degassed for thirty minutes. Grubbs I catalyst (50 mg) was added quickly and the reaction was left to stir until the solution went from purple to red-brown (about six hours). The solution was placed into an SPPS tube, drained, charged with fresh Grubbs I catalyst (50 mg) in 1,2-DCE (4 mL), and returned to the RBF under Ar<sub>(g)</sub> for another six hours. The resin was collected into an SPPS tube and rinsed with DCM until the washes were clear and colorless before adding 3 mL click wash for five-minute washes until the drained liquid came back colorless. The resin was Fmoc-deprotected (3 x 20 minutes) prior to cleavage from resin. The first coupling reaction after the RCM was run overnight to prevent incomplete formation of the amide bond due to Grubbs I catalyst interference.

### **MTT Deprotection**

Prior to an MTT deprotection the resin was washed with DCM (3x). Deprotection occurred with washes of 2% TFA in DCM (5 mL, 9 times) until the drained solvent went

from bright yellow to colorless. The resin was washed with DCM (5x) then NMP (5x), and reswelled in NMP for at least thirty minutes before further manipulation.

#### **Dde Deprotection and Formation of Az4' Group**

The Dde group was removed by incubating with Hydrazine (2 % (v/v) in DMF) (5 mL x 10 minutes) three times before rinsing with DMF (5x). The resin was then reacted with 2-azido acetic acid using the standard hand coupling procedure for a minimum of two hours (**with regular venting for the first thirty minutes of shaking**) and worked up by rinsing with NMP (5x).

#### **Seiber Amide Resin Deprotection to Isolate Fully Protected Peptides**

The resin (3.00 g, 2.01 mmol) was rinsed with DCM (5x), and dried under vacuum for a minimum of one hour. The resin was then washed with TFA (1% (v/v) in DCM) (25 mL x 1 minute once and 10 mL x 1 minute, 10 times) draining into a RBF containing 5mL of DIPEA to neutralize the solution. An additional 1 mL of DIPEA was added to the flask before isolating the crude protein by rotary evaporation and purifying by prep-scale HPLC, identification of pure protected peptides using electrospray ionization mass spectrometry (ESI-MS), and lyophilization of desired fractions.

#### **Resin Cleavage, Global Side-Chain Deprotection, and Purification of Crude Peptide**

The resin was rinsed with DCM (5x) and dried under vacuum for a minimum of one hour. A 20 mL scintillation vial was charged with a stir-bar, resin and cleavage solution (95% TFA, 2.5% TESH, 2.5% MQ H<sub>2</sub>O) before stirring at room temperature for 2 hours. The solution was then filtered into 40 mL of cold ether, vortexed briefly, and stored at 4 °C overnight. The precipitated protein was centrifuged into a pellet (4500 RPM, 10 minutes @ 4 °C) prior to decantation of the supernatant. After dissolving the crude

peptide in DMSO purification occurred by a gradient of 0% to 50% acetonitrile in water with 0.01% (v/v) TFA using an RP-HPLC (Beckman Coulter System Gold 126 Solvent Module and 168 Detector) using a C18 reversed phase semi-preparative column (Phenomenex Luna 10  $\mu\text{m}$ ,  $250 \times 10$  mm), identification of pure peptide using MALDI-TOF mass spectrometry, and lyophilization of desired fractions.

### **Linker Library Synthesis**

The peptide anchors DNB and JXL1 were synthesized on Tentagel resin and cyclized before a split and mix coupling of the linker library on Titan occurred. Each library was split into five wells and couple with Fmoc-D-Pro-OH, Fmoc-Gly-OH, Fmoc-D-Leu-OH, Fmoc-2-aminoisobutyric acid (Fmoc-Aib-OH), or nothing. Recombination of the resin preceded washing and Fmoc deprotection before repeating for an additional four cycles to get a linker library with a variable 0 to 5 amino acids linker length. A final coupling with Boc-Pra-OH preceded global deprotection of side chains with cleavage cocktail (95% TFA, 2.5% TESH, 2.5%  $\text{H}_2\text{O}$ ) and storage under 0.05% sodium azide in water at 4 °C.

### ***Overnight ELISA Procedure***

**Binding Buffer: TBS + 0.05% Tween20 + 0.1% BSA**

**Blocking Buffer: TBS + 5-7% milk + 0.05% Tween20**

**All washing steps were 200 $\mu\text{L}$ /well**

**All incubation steps were 100 $\mu\text{L}$ /well except for the blocking step, which was 200 $\mu\text{L}$ /well**

Prior to incubating with the blank and ligands the 96 well plate was washed with binding buffer (3 x 5 minutes). The wells were then loaded with 100  $\mu\text{L}$  of either 0.5% (v/v) DMSO/ Biotin-PEG<sub>5</sub> linker (2  $\mu\text{M}$  in binding buffer) or biotinylated ligand (2  $\mu\text{M}$  in binding buffer), and the plate was incubated at RT for two hours. The washing step was

repeated before adding blocking solution and the plate was incubated for 12-14 hours in the cold room. After washing the plate in the cold room, each well was charged with either binding buffer or HRP2 protein (20-40 nM in binding buffer), and the plate was left to incubate in the cold room for 7-8 hours. Unbound protein was washed away at room temperature before each well was charged with the anti-GST HRP-conjugated antibody (2000:1 dilution in binding buffer) before incubating at RT for one hour. The plate was then washed with binding buffer (3 x 5 minutes) then with TBS buffer (1 x 5 minutes) before adding developing solution (1:1 TMB Peroxidase Substrate: Peroxidase Substrate Solution B) and agitating until a blue color began to appear. Approximately five minutes after the initial appearance of color the reaction was quenched with 100  $\mu$ L of 1M  $\text{H}_2\text{SO}_{4(\text{aq})}$  at which point the color changed from blue to golden yellow. The absorbance was measured immediately at  $\lambda = 450\text{nm}$ , background corrected, and subsequently plotted in Graphpad.

A similar procedure to the above was utilized to prepare the LVYYY binding curve with the exceptions that the ligand was immobilized at a 200 nM concentration, and the plate was blocked with 3% BSA.

The cocktail ELISA experiments were run using a similar experimental protocol with the exception that the blank used was Biotin-Gly-NHAc and the plate was blocked with 3% BSA.

### ***Room-Temperature ELISA Procedure<sup>1</sup>***

**Binding Buffer: PBS + 0.05% Tween20**

**Blocking Buffer: PBS + 2% BSA+ 0.05% Tween20**

**All washing steps were 200 $\mu$ L/well**



**All incubation steps were 100 $\mu$ L/well except for the blocking step, which was 200 $\mu$ L/well**

Prior to incubating with the biotinylated blank and ligands the 96 well plate was washed with binding buffer (3 x 5minutes). The wells were then loaded with 100  $\mu$ L of either Biotin-PEG<sub>5</sub>-NH<sub>2</sub> (blank) (2  $\mu$ M in binding buffer) or biotinylated ligand (2  $\mu$ M in binding buffer) and the plate incubated with shaking at RT for two hours. The washing step was repeated before adding blocking solution and incubating with shaking for two hours at room temperature. After washing the plate, each well was charged with either binding buffer or a HRP2 protein solution (1-5 nM in binding buffer) and the plate was left to incubate at room temperature for two hours. Unbound protein was washed away at room temperature before each well was charged with the anti-GST HRP Antibody (2000:1 dilution in binding buffer), and the plate was incubated at RT for one hour. The plate was then washed with binding buffer (3 x 5 minutes) and PBS (1 x 5 minutes) before adding 100  $\mu$ L of the developing solution (1:1 TMB Peroxidase Substrate: Peroxidase Substrate Solution B) and agitated until a blue color began to appear. Approximately five to ten minutes after the initial appearance of color the reaction was quenched with 100  $\mu$ L of 1 M H<sub>2</sub>SO<sub>4(aq)</sub> at which point the color changes from blue to golden yellow. The absorbance was measured immediately at  $\lambda$  = 450nm, and subsequently plotted in Graphpad.

### ***Linker Library Screens***

The procedures described herein are a modification of the procedures given in **reference 10** of the main text.

### ***Linker Library Pre-Clear***

Each anchor library was split into three aliquots and pre-cleared separately. Each aliquot was rinsed with MeOH (5x), MQ H<sub>2</sub>O (10x), and swelled in 500  $\mu$ L of binding buffer (0.1% BSA, 0.1% Tween 20 in TBS) @ RT for three hours. The library aliquots were then mixed with 500  $\mu$ L blocking buffer (1% BSA, 0.1% Tween 20 in TBS) and left to shake at 4 °C overnight. After draining and rinsing with 500  $\mu$ L binding buffer the library was incubated for four hours @ RT with 500  $\mu$ L of binding buffer containing 5  $\mu$ M of one of the biotinylated ligands: JXL library DNB, C-term, or N-term; DNB Library JXL1, C-term, or N-term. The beads were then rinsed with binding buffer (3 x 500  $\mu$ L) before incubation with 500  $\mu$ L binding buffer containing a 10,000:1 dilution of the mouse anti-biotin conj. alkaline phosphatase antibody for one hour @ RT. Rinses with binding buffer (3 x 500  $\mu$ L), TBS (1 x 500  $\mu$ L) preceded incubations with a high salt (25 mM Tris•HCl, 10 mM MgCl<sub>2</sub>, 700 mM NaCl, pH = 7.4) TBS buffer @ RT (500  $\mu$ L, 3 x 20 minutes). The beads were then rinsed with BCIP Buffer (3 x 500  $\mu$ L) and dried under vacuum for eight minutes before transferring to a petri dish. The beads were developed with 13  $\mu$ L of BCIP (50 mg/mL in 70% DMF) in 10mL of developing buffer (100 mM Tris•HCl, 150 mM NaCl, 1 mM MgCl<sub>2</sub>, pH=xx) and incubated with the library beads in a 20cm petri dish for ten minutes before adding 26  $\mu$ L of NBT (50 mg/mL in 70% DMF) and incubating for an additional thirty minutes. The development was quenched with 100  $\mu$ L of 6 M HCl<sub>(aq)</sub> and the dark purple beads that represent promiscuous binders to the biotinylated ligand were removed using a 10  $\mu$ L micropipette. Each individual library aliquot was prepared for a subsequent product screen by rinsing with NMP (5x), decolorizing with NMP overnight, rinsing with MeOH (5x), MQ H<sub>2</sub>O (10x), and stripping with 7.5 M Gua • HCl, pH = 2 for thirty minutes @ RT.

### **Linker Library Product Screens**

The aliquots were screened individually against the ligand with which they had been pre-cleared. Each library aliquot was swelled in binding buffer for four hours @ RT before rinsing with binding buffer (10 x 500  $\mu$ L) and incubating with 500  $\mu$ L blocking buffer overnight at 4 °C. The beads were then incubated with 100 nM rHRP2 and 5 $\mu$ M of one of the following: N-term, C-term, JXL1, or DNB ligand in 500  $\mu$ L binding buffer @ RT for six hours. After a quick rinse with binding buffer (500  $\mu$ L), the beads were stripped with 500  $\mu$ L of 7.5M Guan·HCl, pH=2.0 @ RT for one hour. A second overnight block with 500  $\mu$ L of blocking buffer @ 4 °C preceded incubation with a 10,000:1 dilution of mouse anti-biotin-alkaline phosphatase conj. Ab in 500  $\mu$ L binding buffer for one hour @ RT. The beads were washed with binding buffer (3 x 500  $\mu$ L), TBS, pH = 7.4 (3 x 500  $\mu$ L) before incubating with high salt buffer (4 x 15 minutes). Finally they were rinsed with TBS (3 x 500  $\mu$ L) and BCIP Buffer (pH = 9, 3 x 500  $\mu$ L), and dried under vacuum for one minute before development in a 20 cm petri dish with 26  $\mu$ L BCIP, 13  $\mu$ L NBT in 15 mL of BCIP buffer for thirty minutes. The development reaction was quenched with 100  $\mu$ L 6 M HCl<sub>(aq)</sub>, and the dark purple beads were picked with a 10  $\mu$ L micropipette. The beads were placed individually into Spinex tubes and washed with 250  $\mu$ L 7.5 M Guan•HCl, pH = 2.0, MQ H<sub>2</sub>O (5 x 250  $\mu$ L), and stored in NaN<sub>3</sub> (0.05 % (<sup>m</sup>/<sub>v</sub>)) in MQ H<sub>2</sub>O at 4 °C until ready to be sequenced.

### **References**

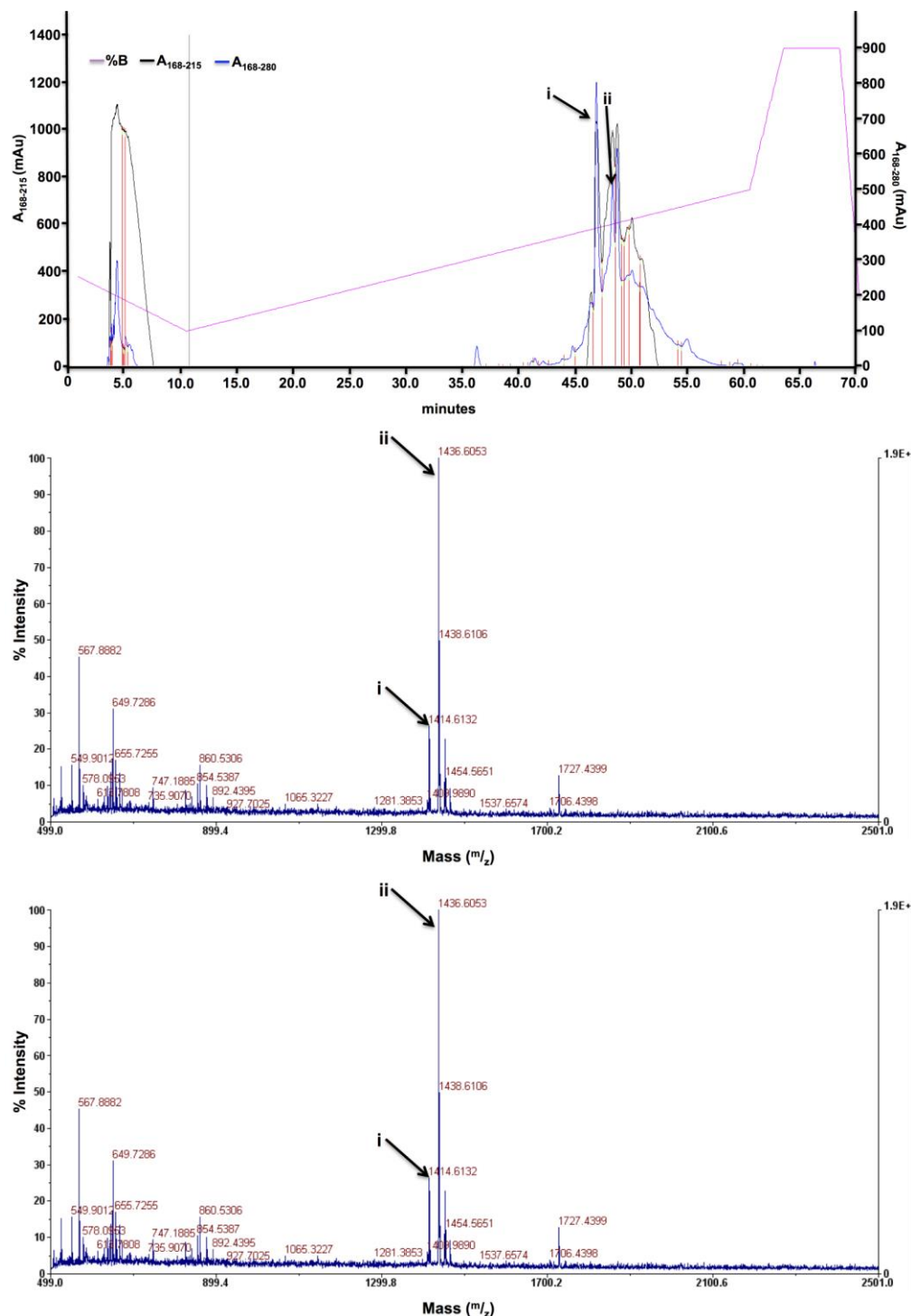
- (1) J. N. Waitumbi, J. Gerlach, I. Afonina, S. B. Anyona, J. N. Koros, J. Siangla, I. Ankoudinova, M. Singhal, K. Watts, M. E. Polhemus, N. M. Vermeulen, W. Mahoney, M. Steele and G. J. Domingo, *Trop. Med. Int. Heal.*, 2011, **16**, 786–793.

## **Appendix 3B**

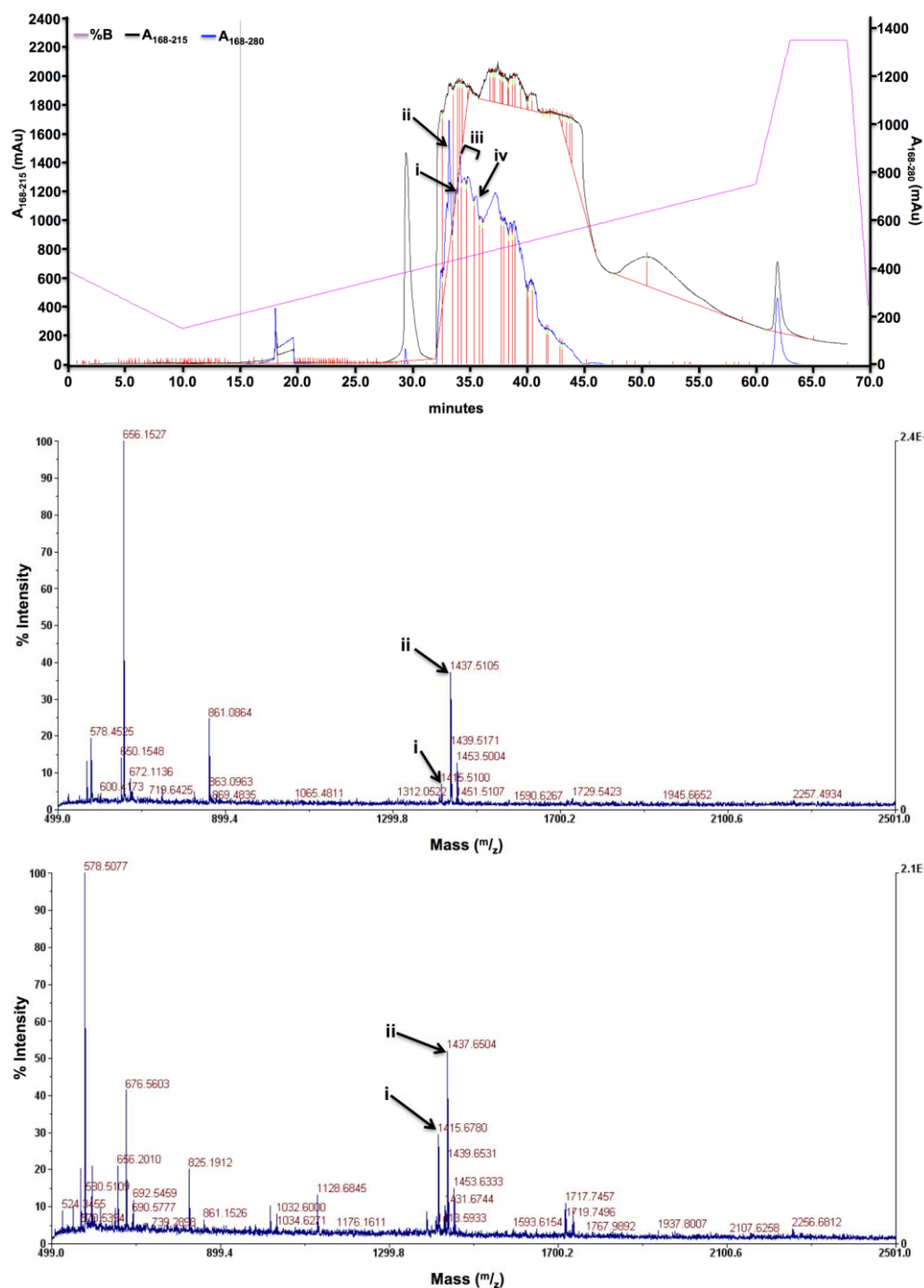
### **Supplemental Figures**

#### **Table of Contents**

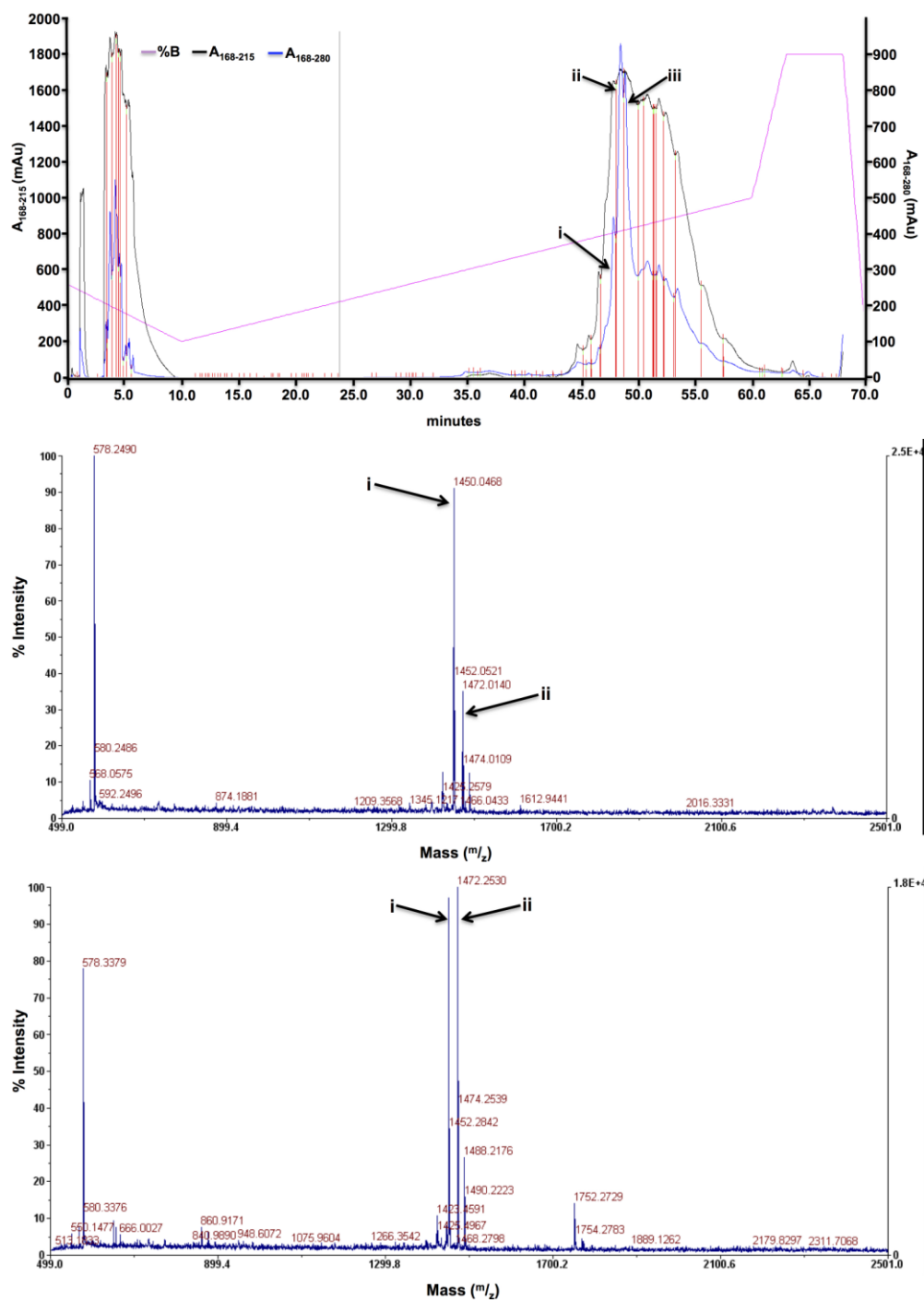
Figure A3B.1-37 Characterization spectra	127
Figure A3B.38-Structure of the spontaneously formed heterodimer in the click reaction.	164



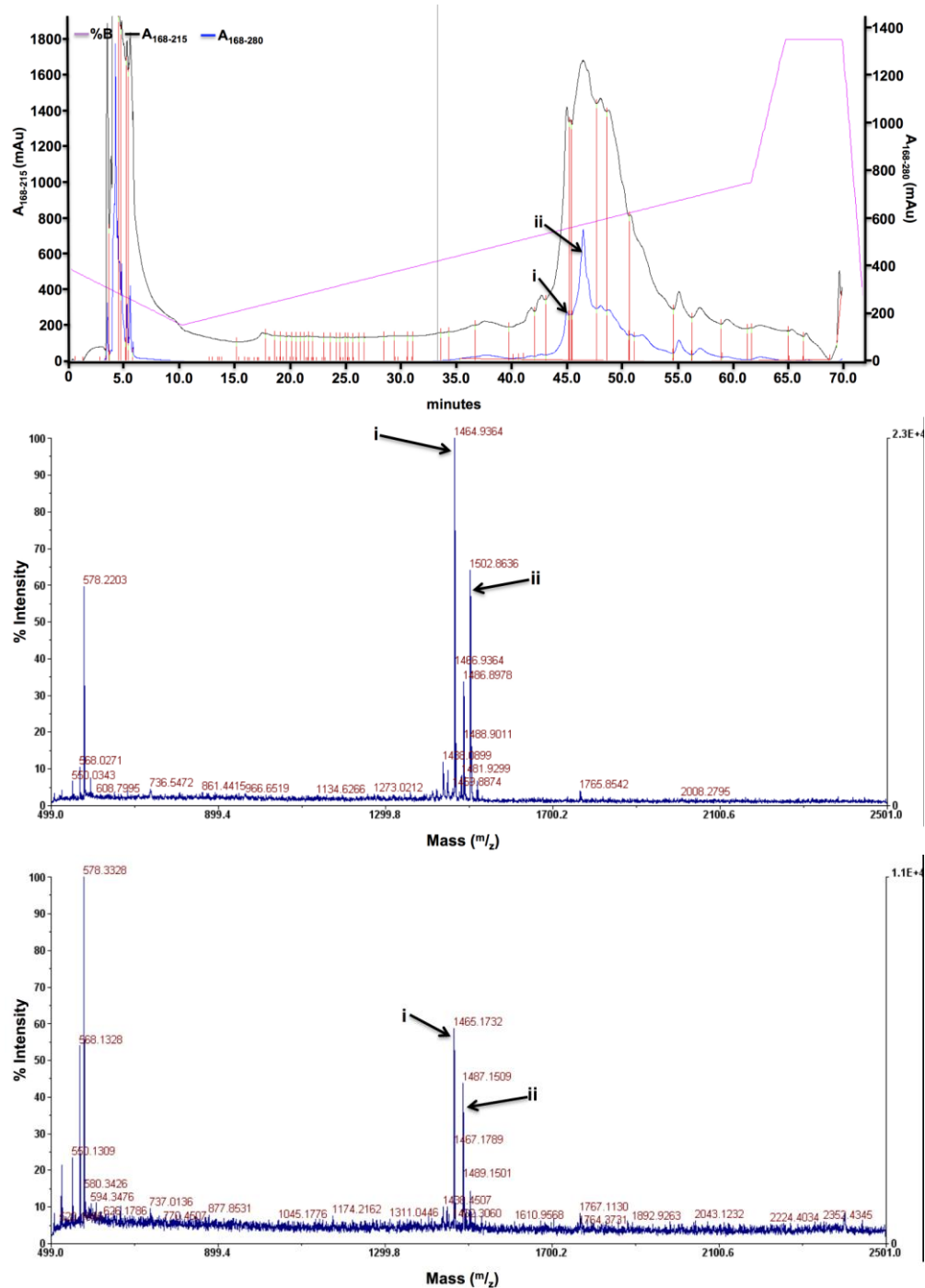
**Figure A3B.1. Characterization spectra for ligand hL1A1.** Ligand hL1A1:  $C_{65}H_{106}N_{16}O_{17}S$ . (Top) The HPLC trace for ligand hL1A1. The arrows correspond to the fractions with the correct mass. (Middle) The MALDI-TOF spectra for the first fraction. The peak labeled “i” corresponds to the  $[M+H]^+$  mass, expected  $[M+H]^+ = 1416.677$ , and the peak labeled “ii” corresponds to the  $[M+Na]^+$  mass, expected  $[M+Na]^+ = 1438.762$ . (Bottom) The MALDI-TOF spectra for the second fraction. The peak labeled “i” corresponds to the  $[M+H]^+$  mass, and the peak labeled “ii” corresponds to the  $[M+Na]^+$  mass.



**Figure A3B.2. Representative characterization spectra for ligand hL1A2.** Ligand hL1A2  $C_{65}H_{106}N_{16}O_{17}S$ . (Top) The HPLC trace for ligand hL1A2. The indicated regions correspond to fractions 1-4 that had the correct mass. (Middle and Bottom) The MALDI spectra for fractions 1 and 2 respectively. The peaks labeled “i” correspond to the  $[M+H]$  mass, expected  $[M+H] = 1416.677$ , and the peaks labeled “ii” correspond to the  $[M+Na]$  mass, expected  $[M+Na] = 1438.762$ .

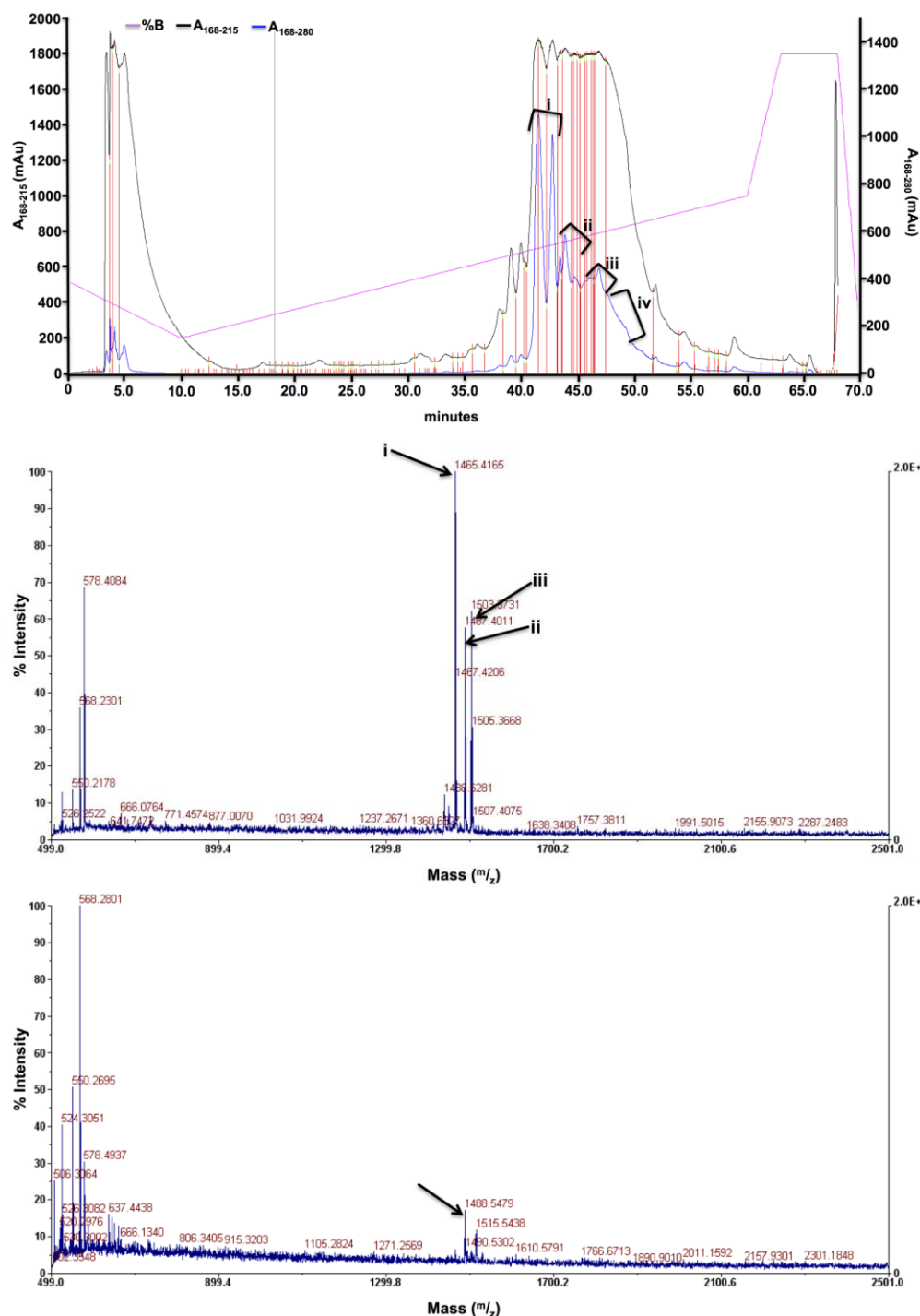


**Figure A3B.3. Representative characterization spectra for ligand hL1A3.** Ligand hL1A3  $C_{69}H_{107}N_{15}O_{17}S$ . (Top) The HPLC trace for ligand hL1A3. The indicated regions correspond to fractions 1-3 that had the correct mass. (Middle and Bottom) The MALDI spectra for fractions 1 and 2 respectively. The peaks labeled “i” correspond to the  $[M+H]^+$  mass, expected  $[M+H] = 1450.777$ , and the peaks labeled “ii” correspond to the  $[M+Na]^+$  mass, expected  $[M+Na] = 1472.759$ .

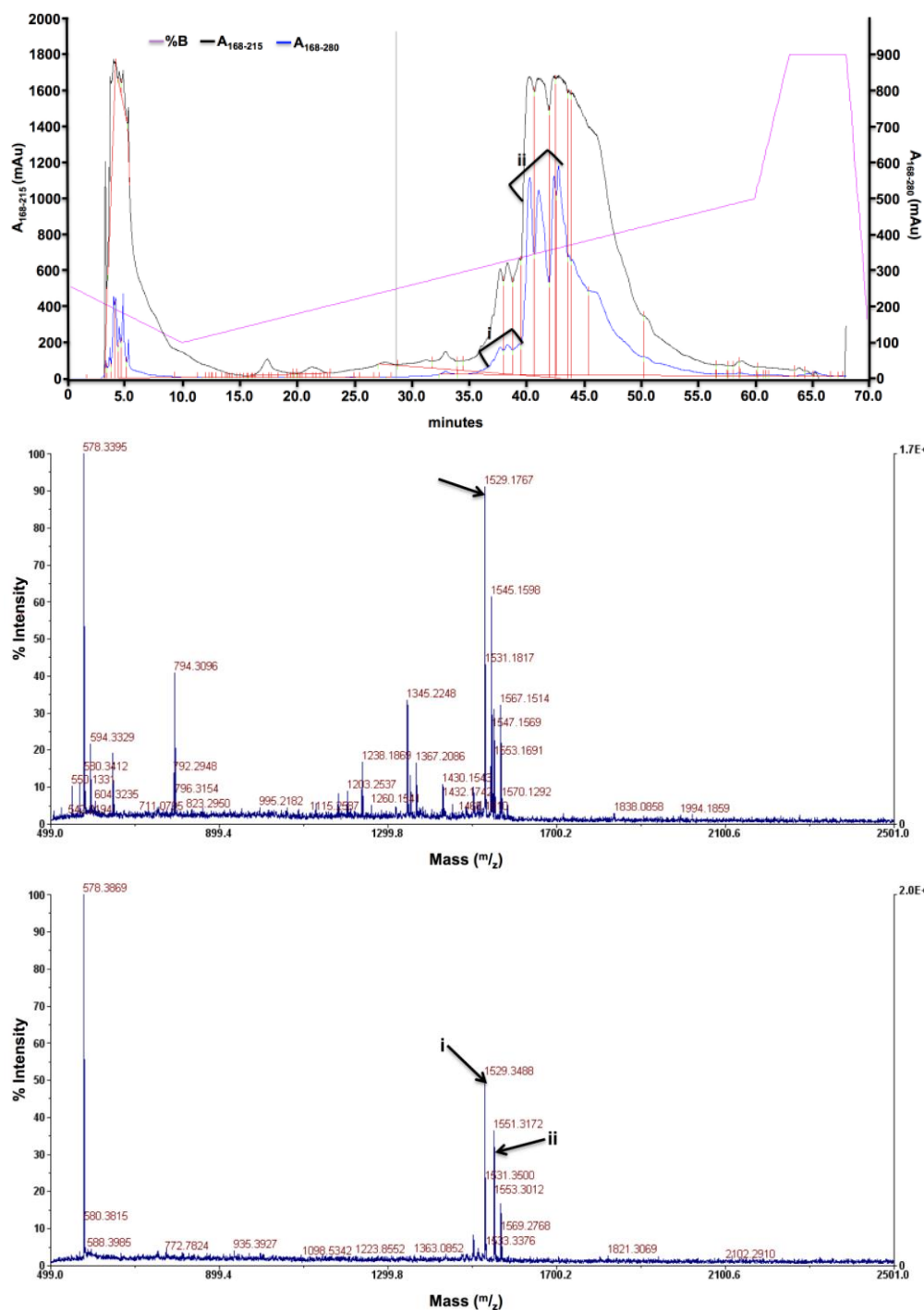


**Figure A3B.4. Characterization spectra for ligand **hL1A4**.** Ligand **hL1A4**  $C_{68}H_{104}N_{16}O_{18}S$ . (Top) The HPLC trace for ligand **hL1A4**. The indicated regions correspond to the two fractions that had the correct mass. (Middle and Bottom) The MALDI spectra for fractions 1 and 2 respectively. The peaks labeled “i” correspond to the  $[M+H]^+$  peak, expected  $[M+H] = 1465.751$ , and the peaks labeled “ii” correspond to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1487.733$ .



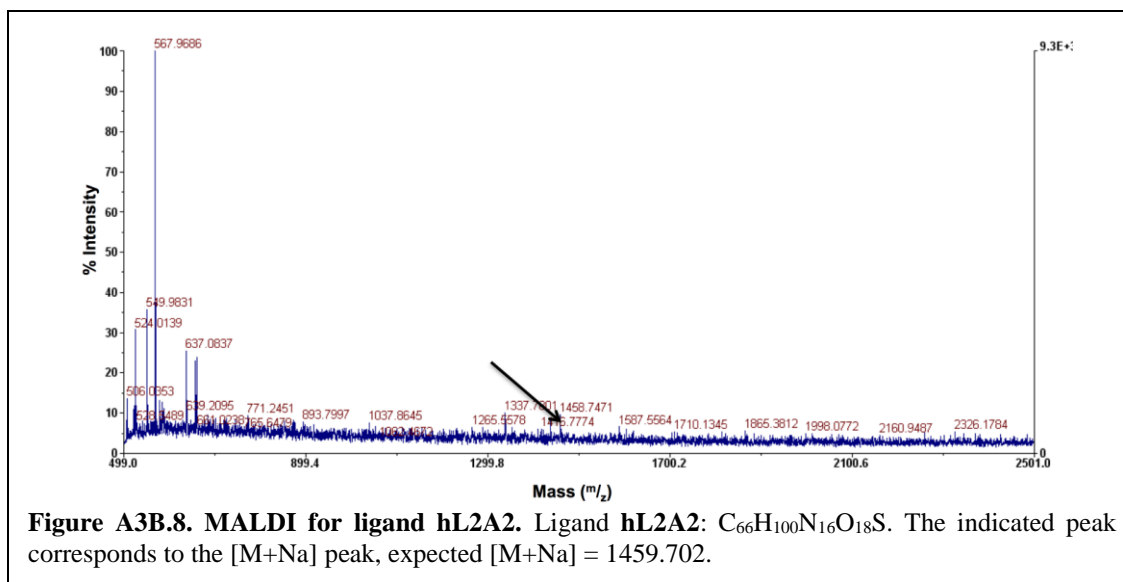


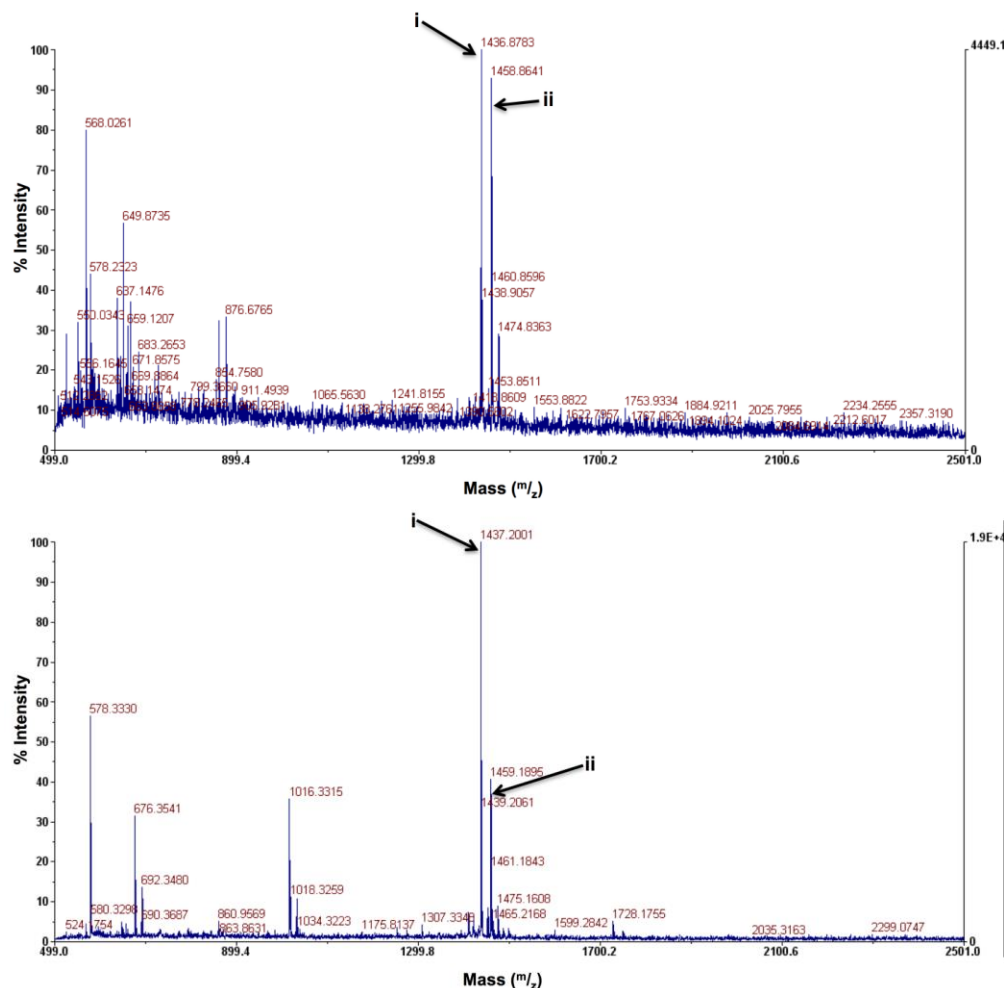
**Figure A3B.5. Representative characterization spectra for ligand hL1A5.** Ligand hL1A5: C<sub>68</sub>H<sub>104</sub>N<sub>16</sub>O<sub>18</sub>S. (Top) The HPLC trace for ligand hL1A5. The indicated regions correspond to the four fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The peaks labeled “i” correspond to the [M+H]<sup>+</sup> peak, expected [M+H]<sup>+</sup> = 1465.751, the peaks labeled “ii” correspond to the [M+Na]<sup>+</sup> peak, expected [M+Na]<sup>+</sup> = 1487.733, and the peak labeled “iii” corresponds to the [M+K]<sup>+</sup> peak, expected [M+K]<sup>+</sup> = 1503.707. (Bottom) The MALDI spectra for fraction 2. The indicated peak corresponds to the [M+Na]<sup>+</sup> peak.



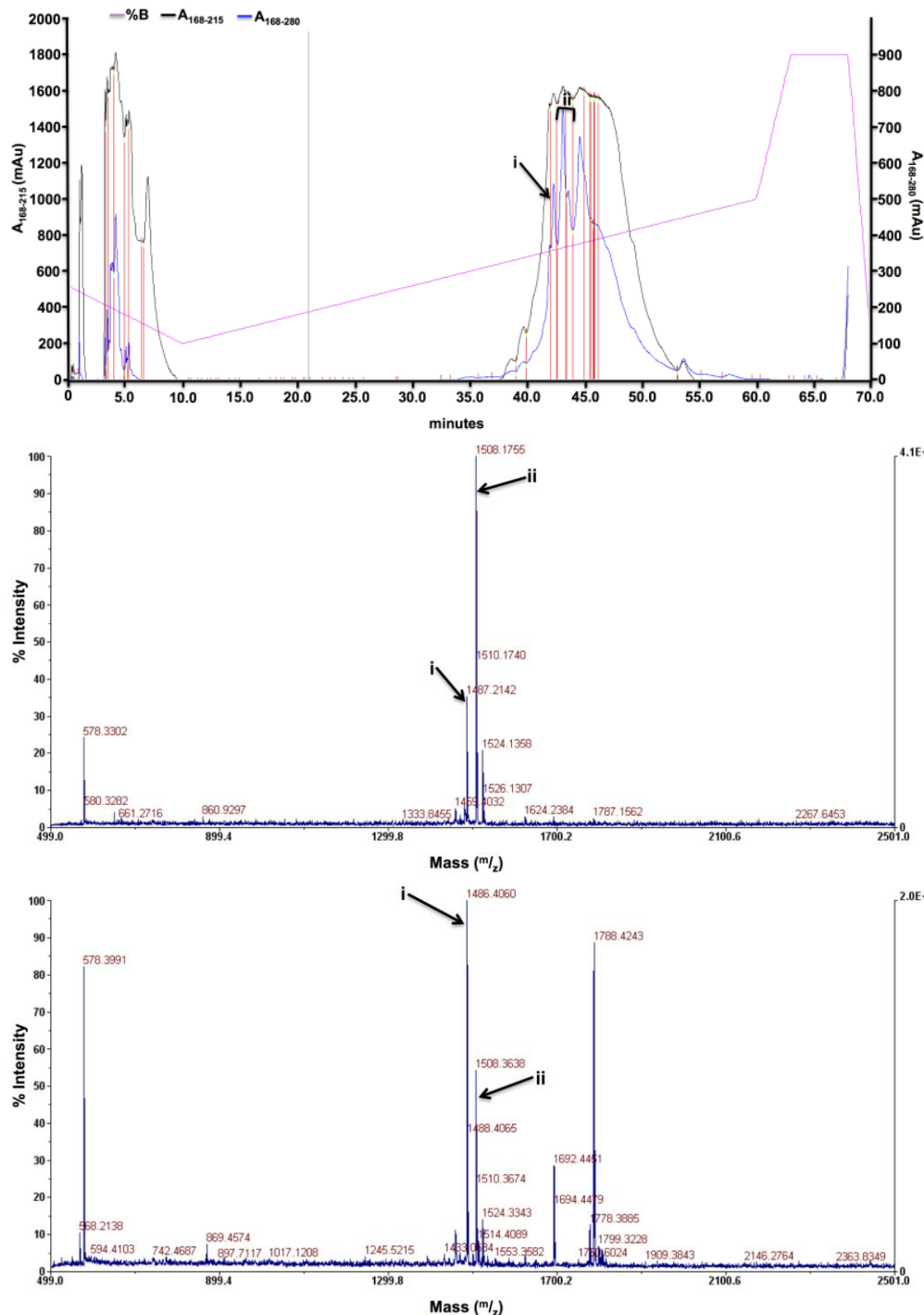
**Figure A3B.6. Characterization spectra for ligand **hL2**.** Ligand **hL2**  $C_{72}H_{104}N_{16}O_{19}S$ . (Top) The HPLC trace for ligand **hL2**. The indicated regions correspond to the two fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The indicated peak corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1530.765$ . (Bottom) The MALDI spectra for fraction 2. The peak labeled “i” corresponds to the  $[M+H]$  peak, and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na] = 1552.747$ .

**Figure A3B.7. Representative characterization spectra for ligand hL2A1.** Ligand **hL2A1**  $C_{66}H_{100}N_{16}O_{18}S$ . (Top) The HPLC trace for ligand **hL2A1**. The indicated regions correspond to the three fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H] = 1437.720$ , and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1459.702$ . (Bottom) The MALDI spectra for fraction 2. The indicated peak corresponds to the  $[M+Na]^+$  peak.

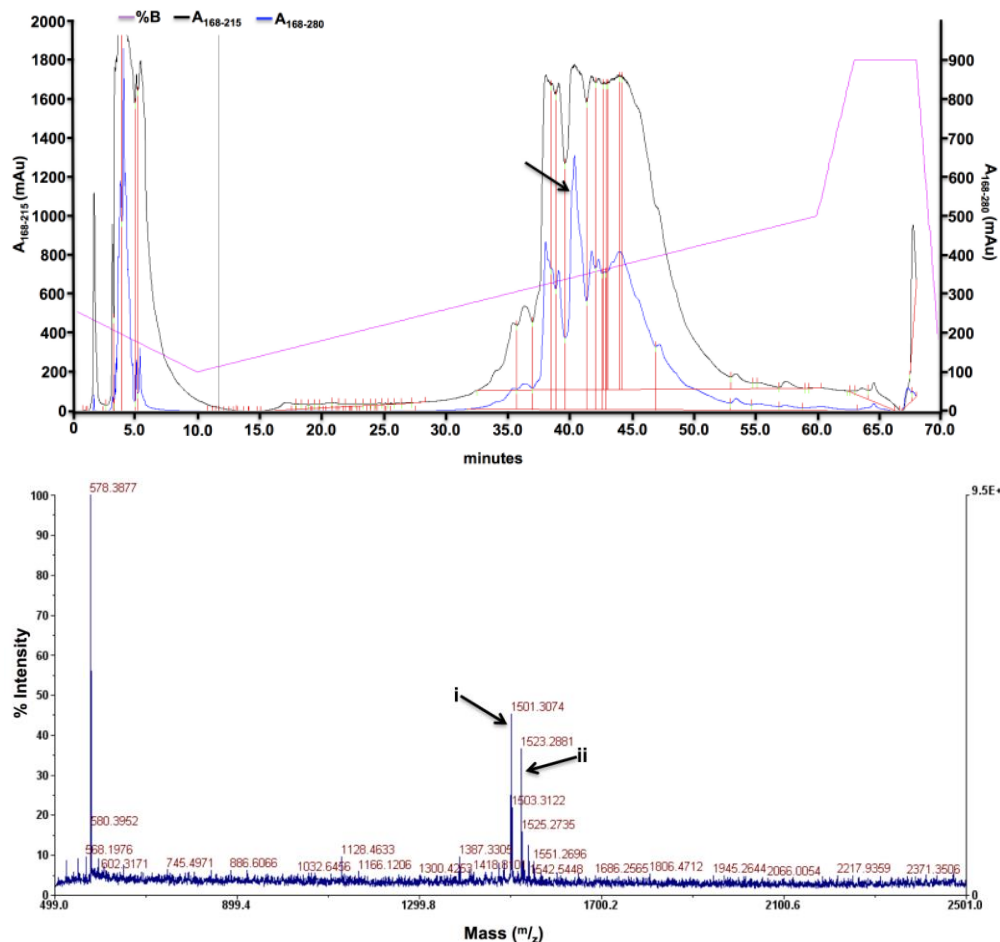




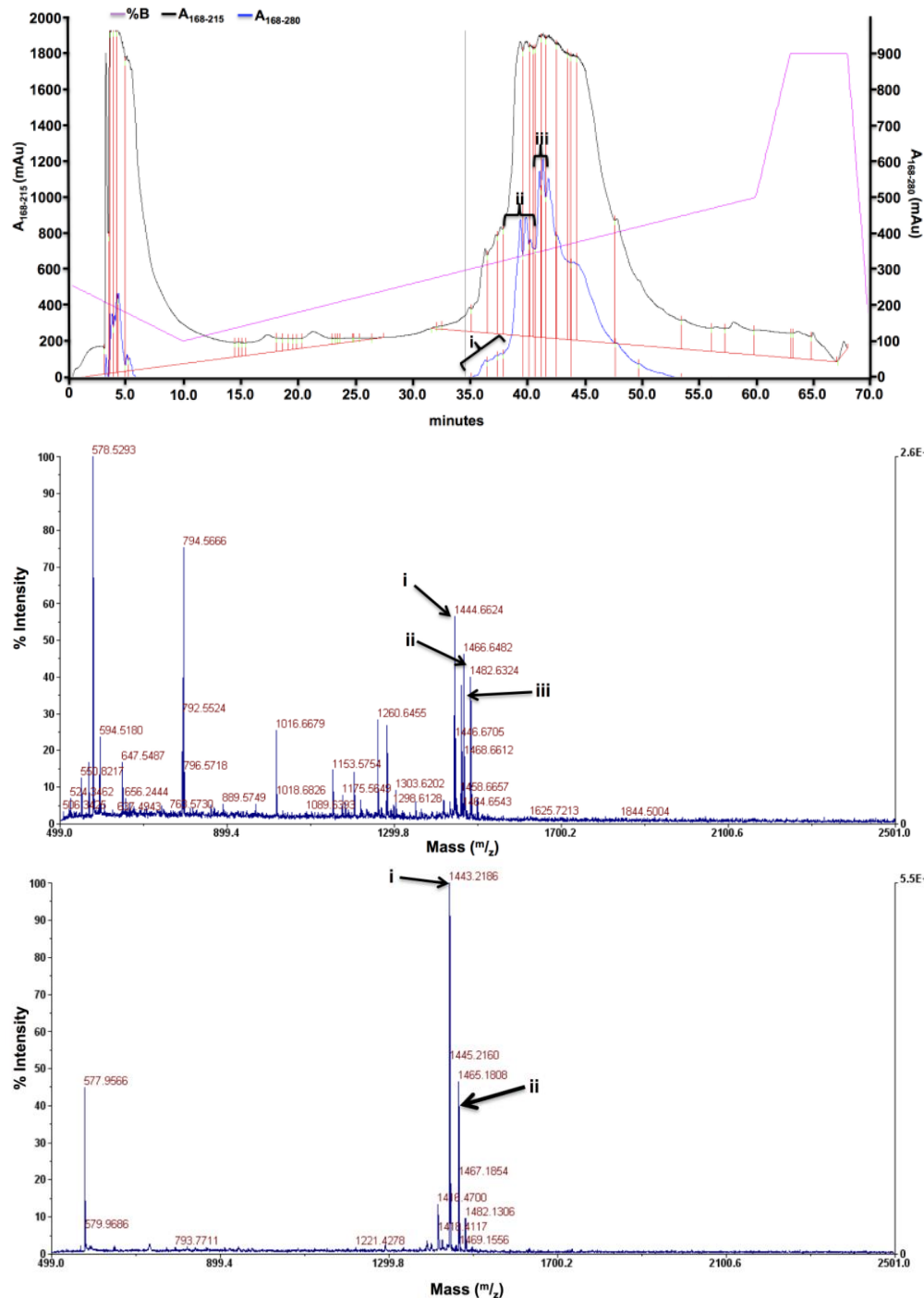
**Figure A3B.9. MALDI spectra for ligand hL2A3.** Ligand hL2A3:  $C_{66}H_{100}N_{16}O_{18}S$ . The MALDI spectra for fractions one and two of L2A3 respectively. The peaks labeled “i” correspond to the [M+H] peak, expected [M+H] = 1437.720, and the peaks labeled “ii” correspond to the [M+Na] peak, expected [M+Na] = 1459.702.



**Figure A3B.10. Characterization spectra for ligand hL2A4.** Ligand **hL2A4**:  $C_{71}H_{103}N_{15}O_{18}S$ . (Top) The HPLC trace for ligand **hL2A4**. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The peaks labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H]^+=1486.740$ , and the peaks labeled “ii” correspond to the  $[M+Na]^+$  peak, expected  $[M+Na]^+=1508.722$ .

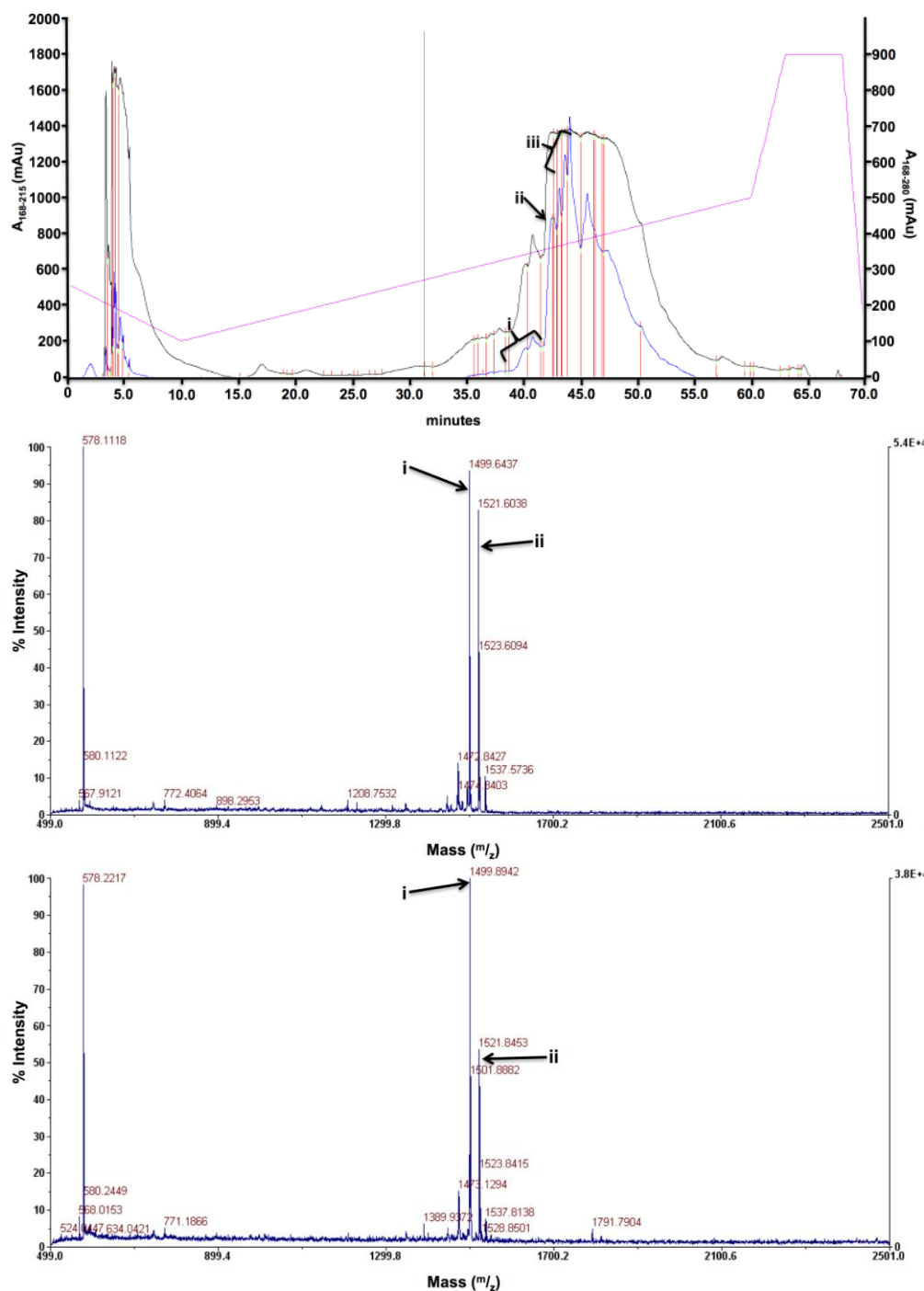


**Figure A3B.11. Characterization spectra for ligand hL2A5.** Ligand **hL2A5**  $C_{70}H_{100}N_{16}O_{19}S$ . (Top) The HPLC trace for ligand **hL2A5**. The indicated regions correspond to the fraction that had the correct mass. (Bottom) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H]=1501.715$ , and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na]=1523.697$ .

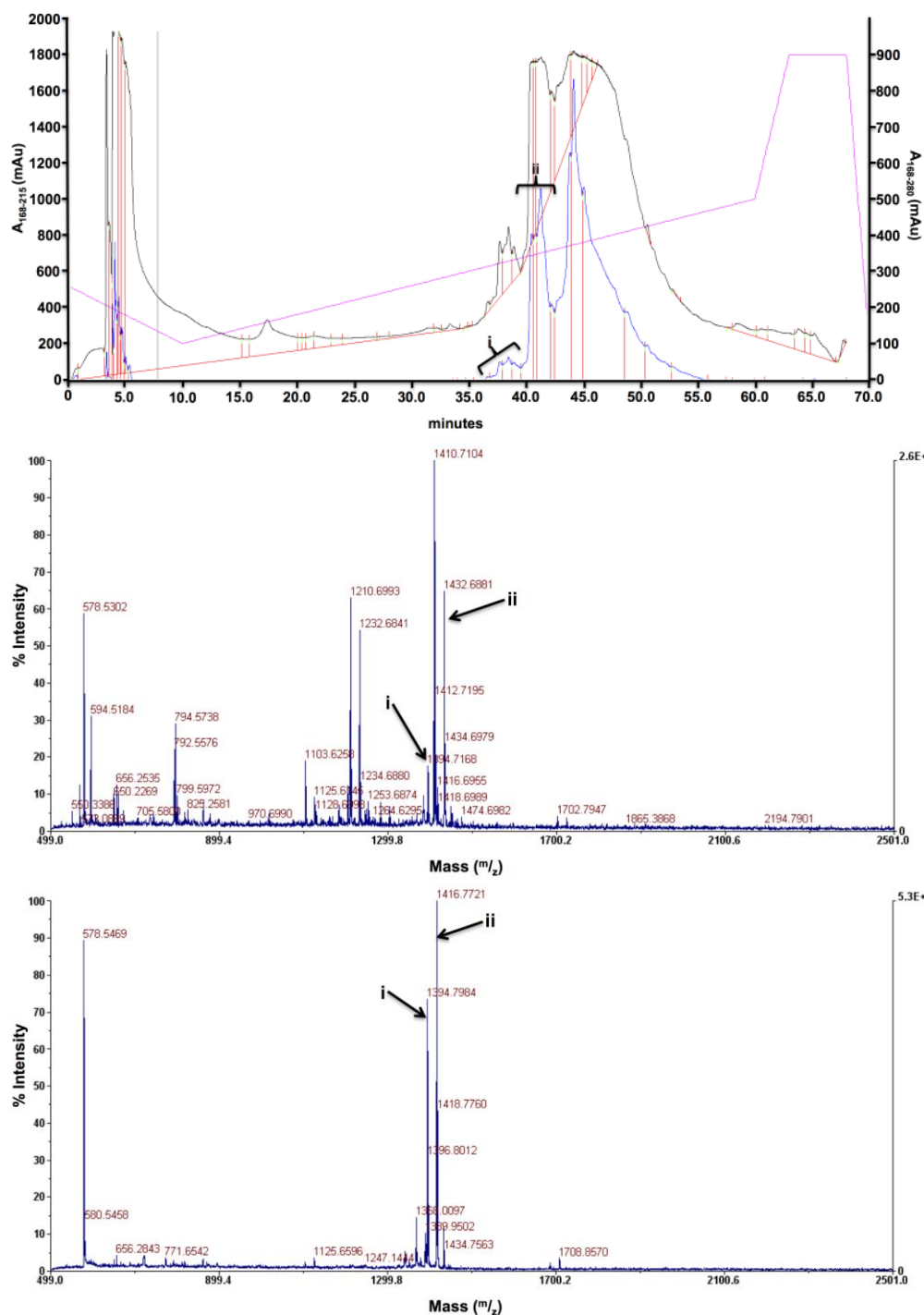


**Figure A3B.12. Representative characterization spectra for ligand hL3A1.** Ligand hL3A1:  $C_{68}H_{97}N_{15}O_{18}S$ . (Top) The HPLC trace for ligand hL3A1. The indicated regions correspond to the three fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The peaks labeled “i” correspond to the  $[M+H]^+$  peak, expected  $[M+H] = 1444.693$ , the peaks labeled “ii” correspond to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1466.675$ , and the peak labeled “iii” corresponds to the  $[M+K]^+$  peak, expected  $[M+K] = 1482.649$ . (Bottom) The MALDI spectra for fraction 2. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak.

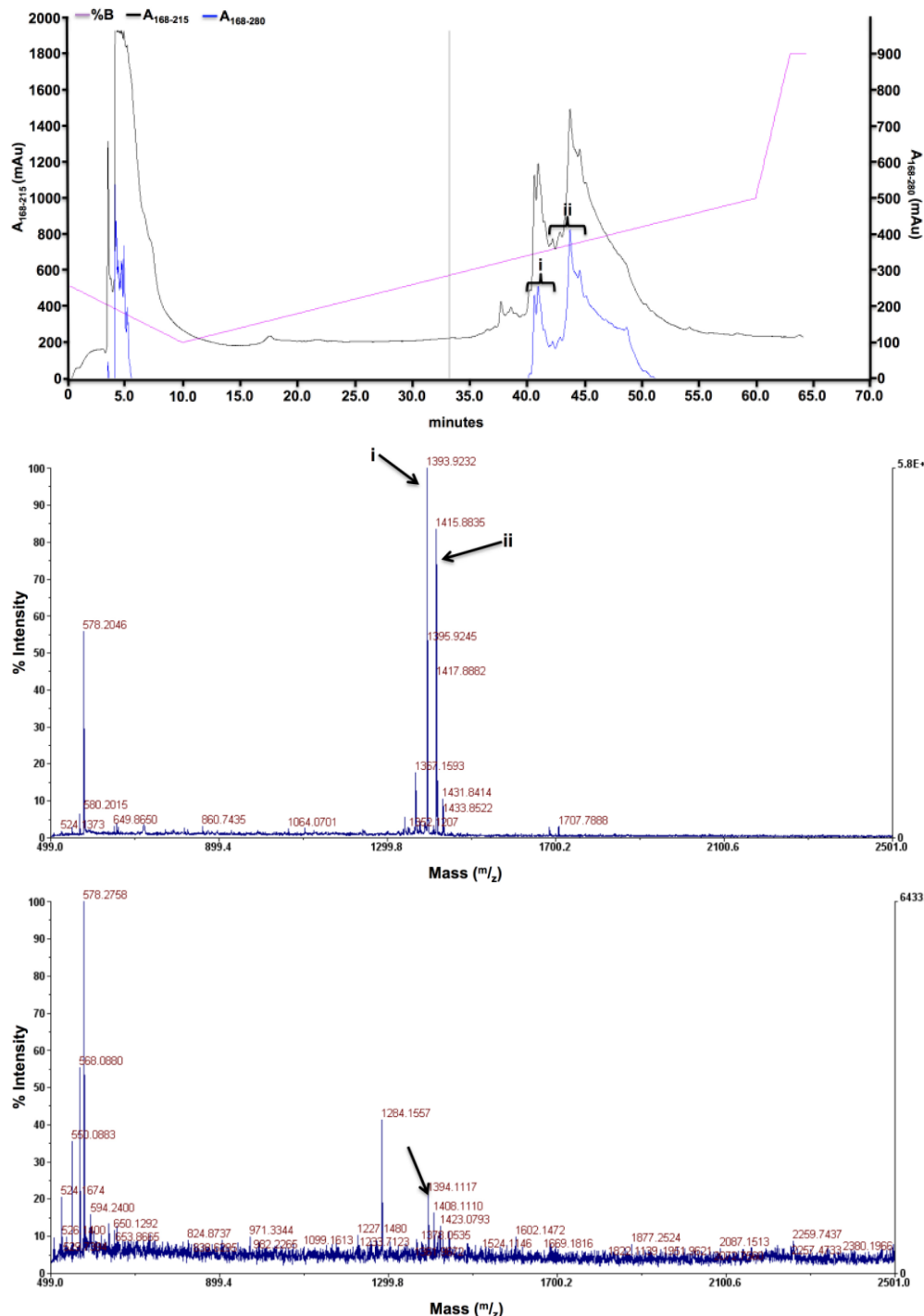




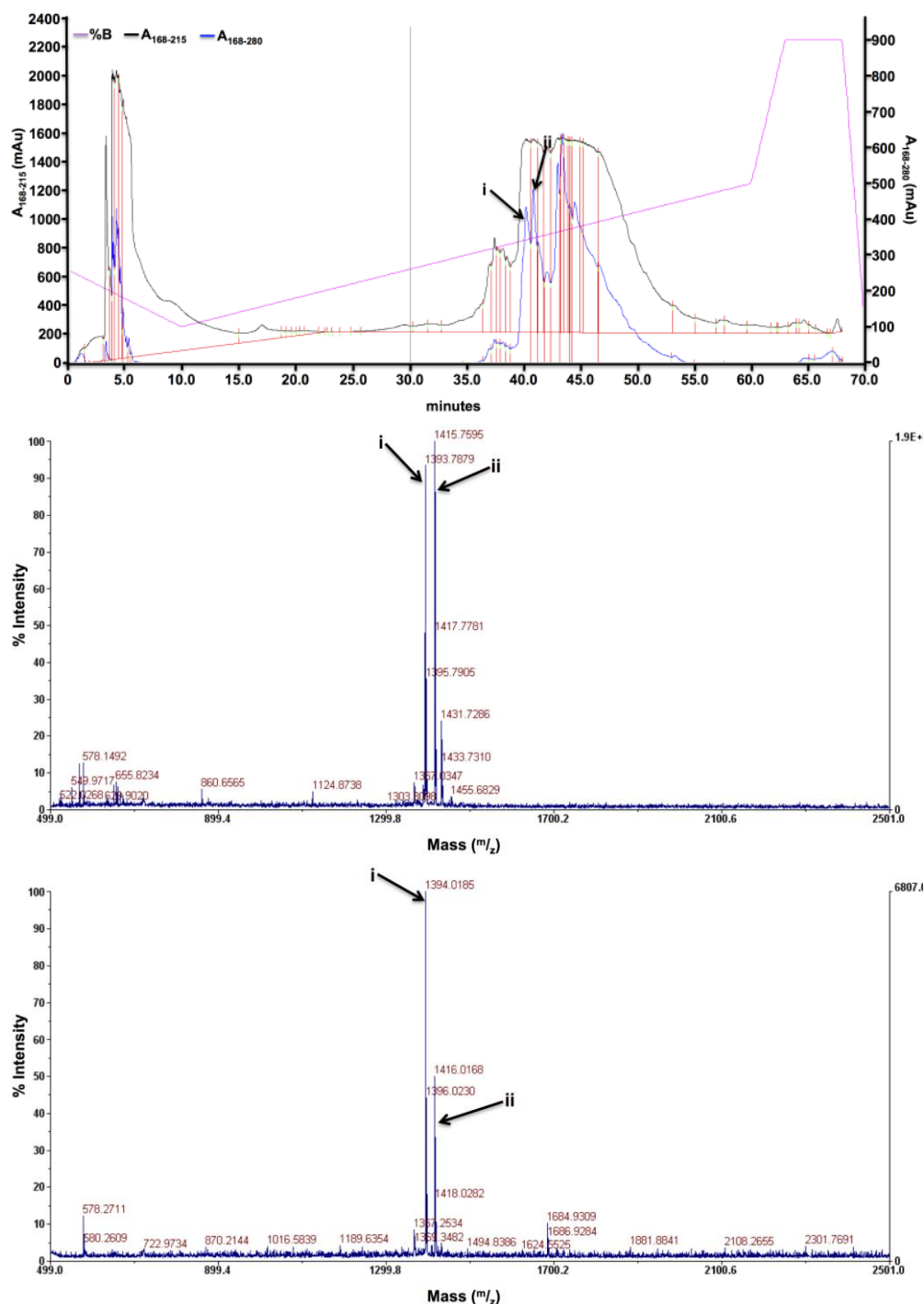
**Figure A3B.13. Representative characterization spectra for ligand hL3A2.** Ligand hL3A2:  $C_{72}H_{105}N_{15}O_{18}S_r$ . (Top) The HPLC trace for ligand hL3A2. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 2 and fraction 3 respectively. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H]^+ = 1500.756$  and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak, expected  $[M+Na]^+ = 1522.738$ .



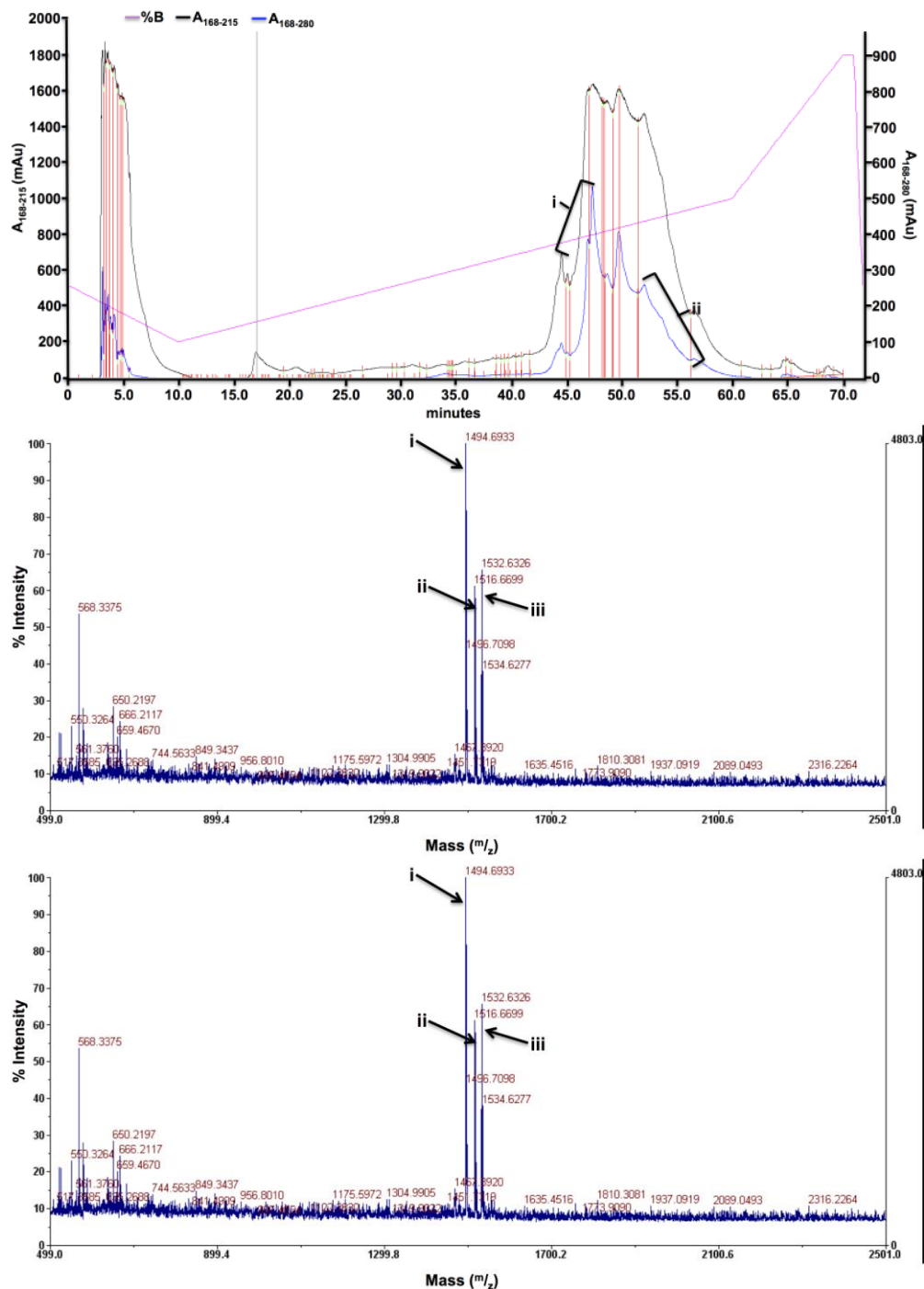
**Figure A3B.14. Characterization spectra for ligand hL3A3.** Ligand hL3A3:  $C_{65}H_{99}N_{15}O_{17}S$ . (Top) The HPLC trace for ligand hL3A3. The indicated regions correspond to the two fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1394.714$ , the peak labeled “ii” corresponds to the  $[M+K]$  peak, expected  $[M+K] = 1432.670$ . (Bottom) The MALDI spectra for fraction 2. The peak labeled “i” corresponds to the  $[M+H]$  peak, and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na] = 1416.696$ .



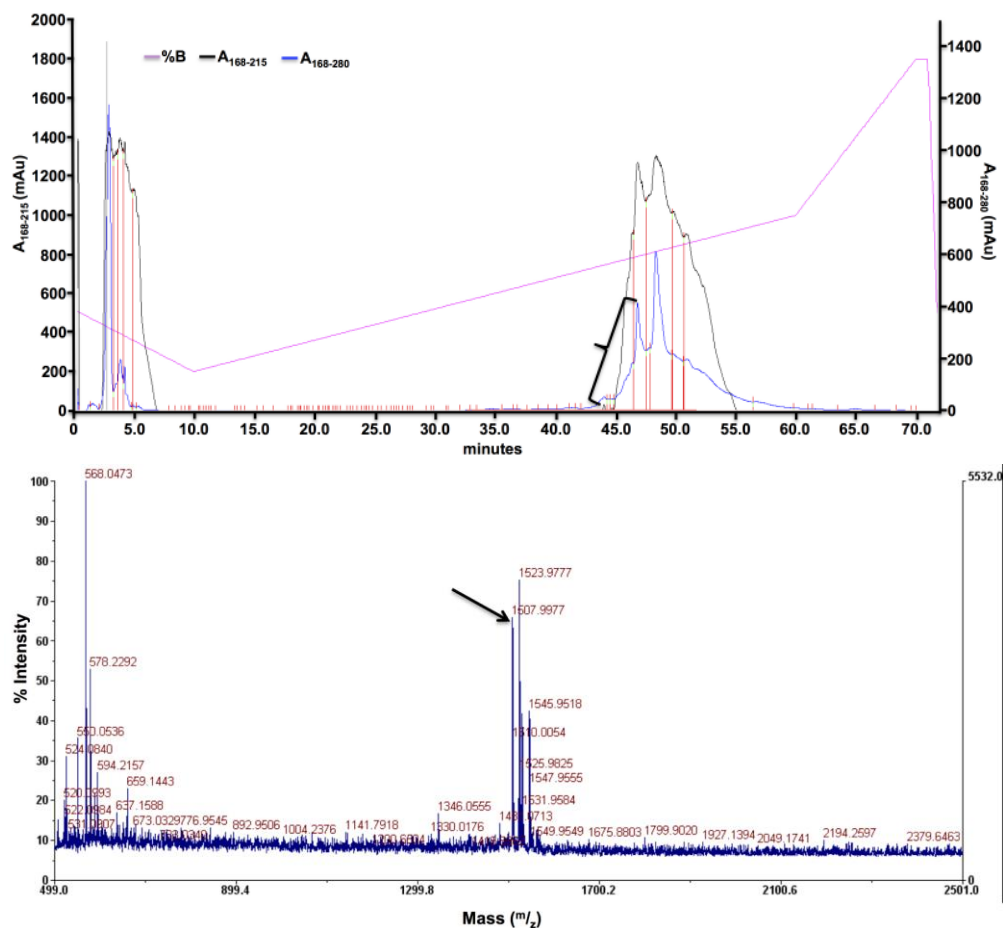
**Figure A3B.15. Characterization spectra for ligand hL3A4.** Ligand hL3A4: C<sub>65</sub>H<sub>99</sub>N<sub>15</sub>O<sub>17</sub>S. (Top) The HPLC trace for ligand hL3A4. The indicated regions correspond to the two fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the [M+H] mass, expected [M+H] = 1394.714, and the peak labeled “ii” corresponds to the [M+Na] mass, expected [M+Na] = 1416.696. (Bottom) The MALDI spectra for fraction 2. The indicated peak corresponds to the [M+H] mass.



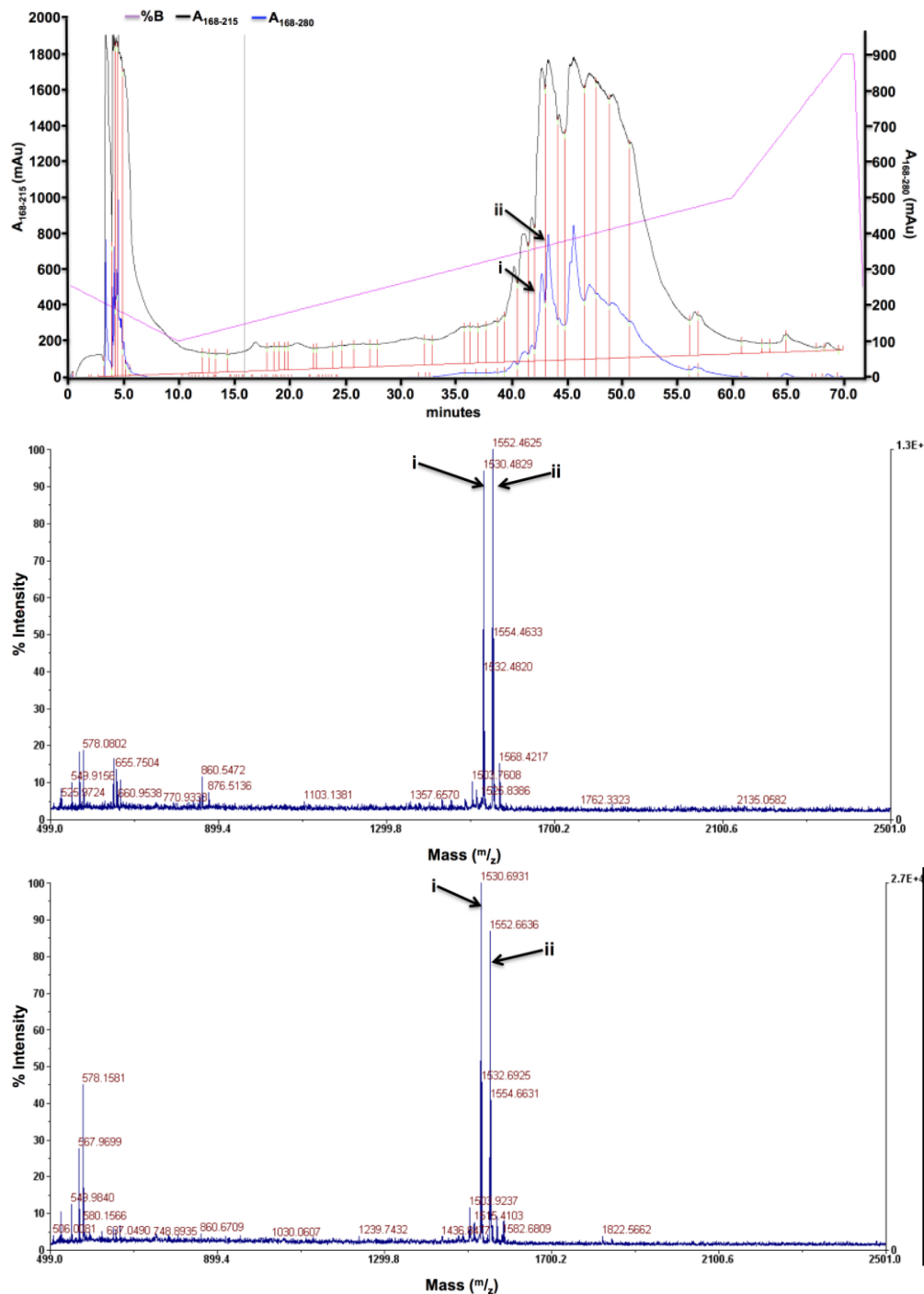
**Figure A3B. 16. Characterization spectra for ligand hL3A5.** Ligand hL3A5:  $C_{65}H_{99}N_{15}O_{17}S$ . (Top) The HPLC trace for ligand hL3A5. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H]^+ = 1394.714$ , and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak, expected  $[M+Na]^+ = 1416.696$ .



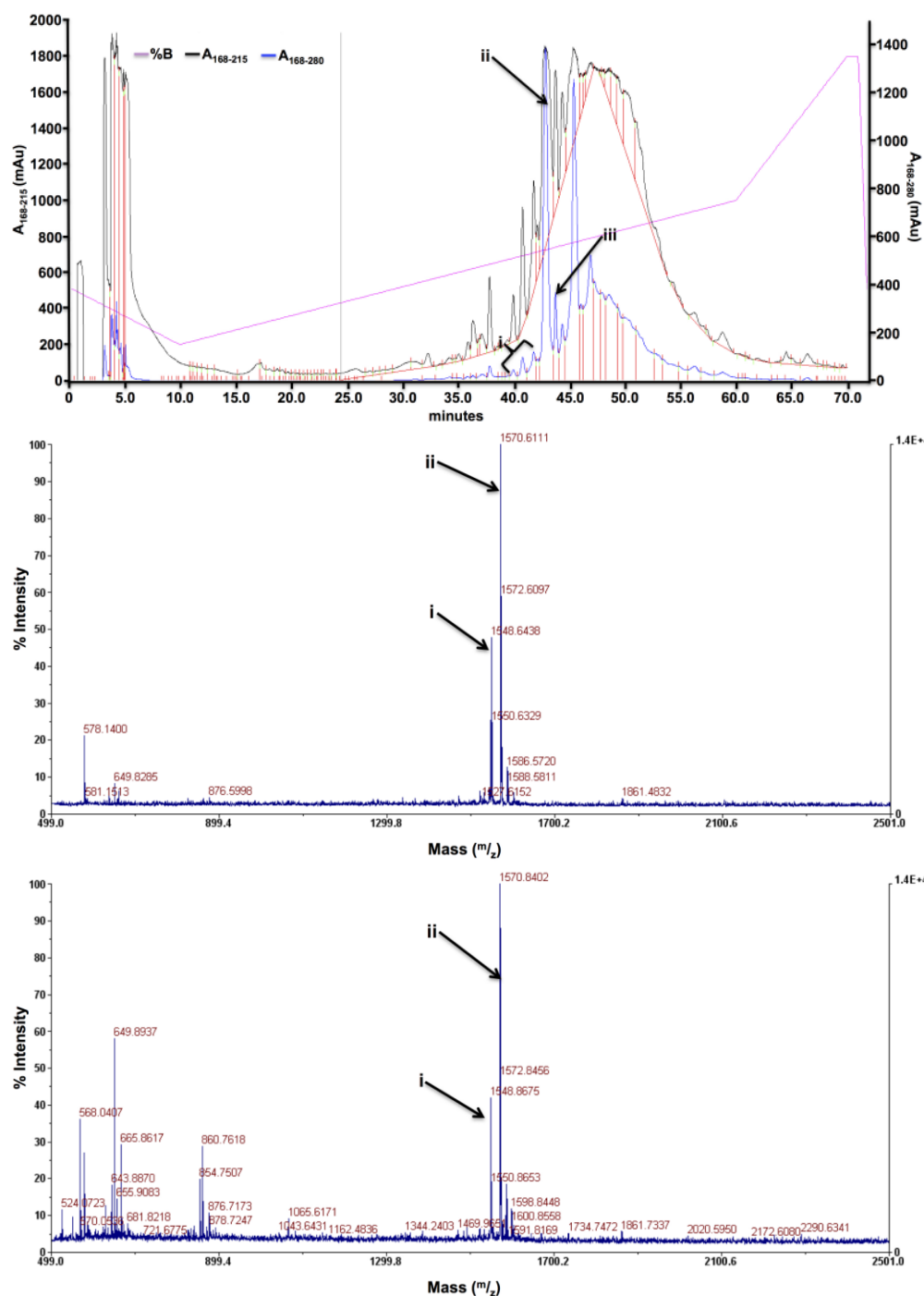
**Figure A3B.17. Characterization spectra for ligand hL1D3.** Ligand **hL1D3**:  $C_{70}H_{107}N_{15}O_{19}S$ . (Top) The HPLC trace for ligand **hL1D3**. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The peaks labeled “i” correspond to the  $[M+H]^+$  peak, expected  $[M+H] = 1494.767$ , the peaks labeled “ii” correspond to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1516.749$ , and the peaks labeled “iii” correspond to the  $[M+K]^+$  peak, expected  $[M+K] = 1532.723$ .



**Figure A3B.18. Characterization spectra for hL1E3.** Ligand **hL1E3**:  $C_{71}H_{109}N_{15}O_{19}S$ . (Top) The HPLC trace for ligand **hL1E3**. The indicated regions correspond to the fraction that had the correct mass. (Bottom) The MALDI spectra for ligand **hL1E3**. The indicated peak corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1509.787$ .

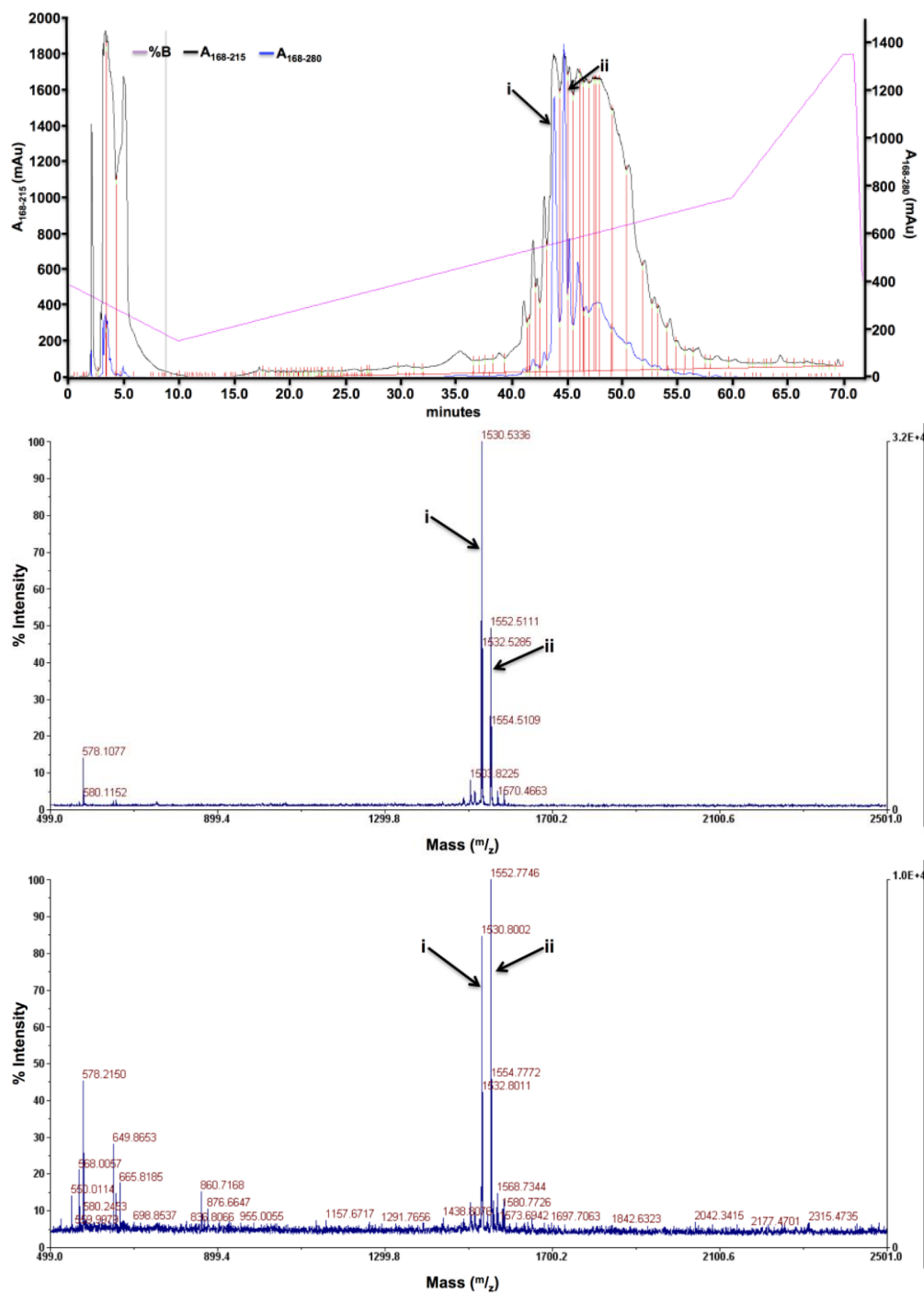


**Figure A3B.19. Characterization spectra for ligand **hL2fF1**.** Ligand **hL2fF1**:  $C_{72}H_{103}FN_{16}O_{18}S$ . (Top) The HPLC trace for ligand **hL2fF1**. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H] = 1531.742$ , and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1553.724$ .

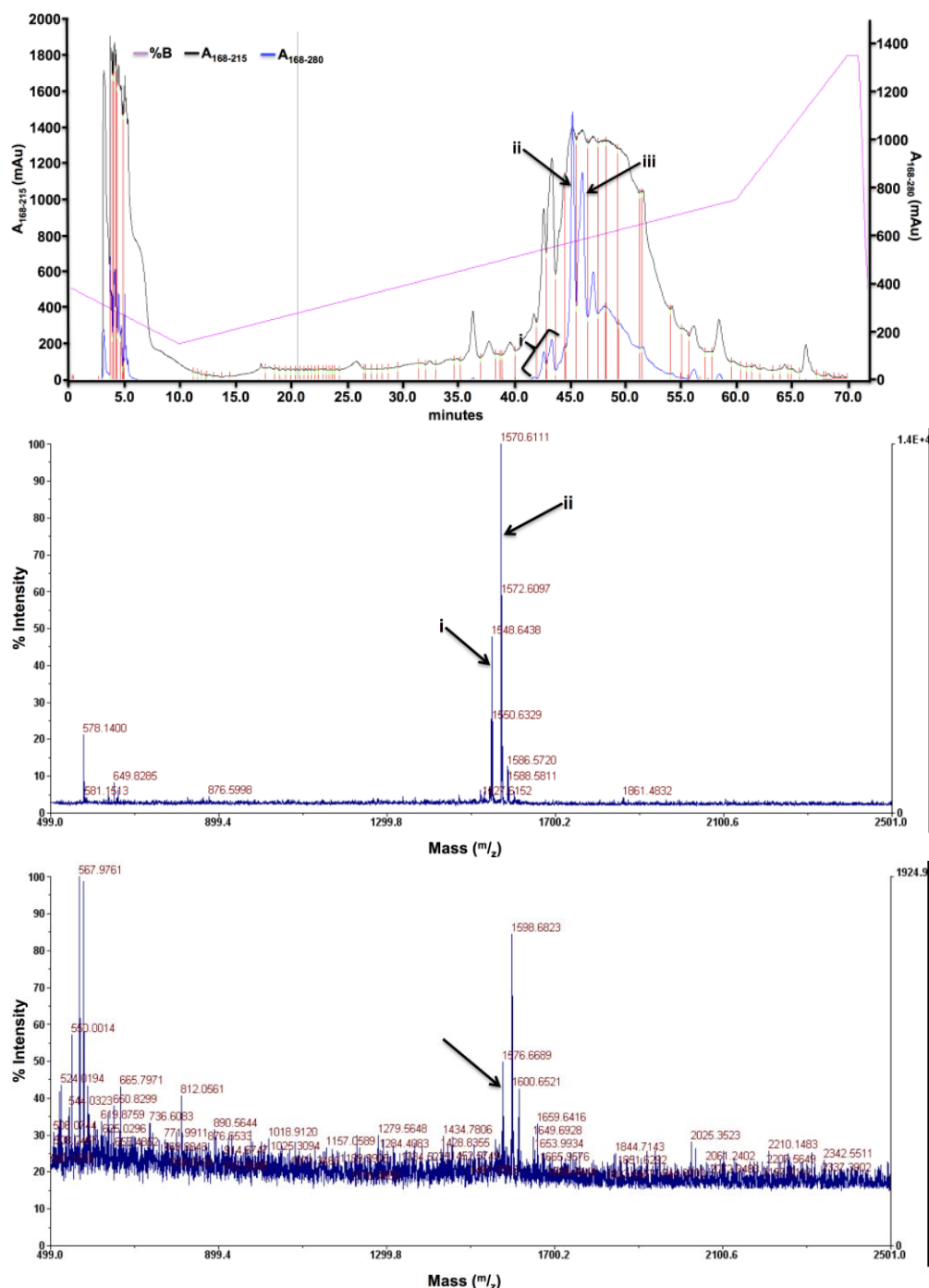


**Figure A3B.20. Representative characterization spectra for ligand hL2f\*F1.** Ligand hL2f\*F1  $C_{72}H_{102}F_2N_{16}O_{18}S$ . (Top) The HPLC trace for ligand hL2f\*F1. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 2 and fraction 3 respectively. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H]=1549.733$ , and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na]=1571.714$ .

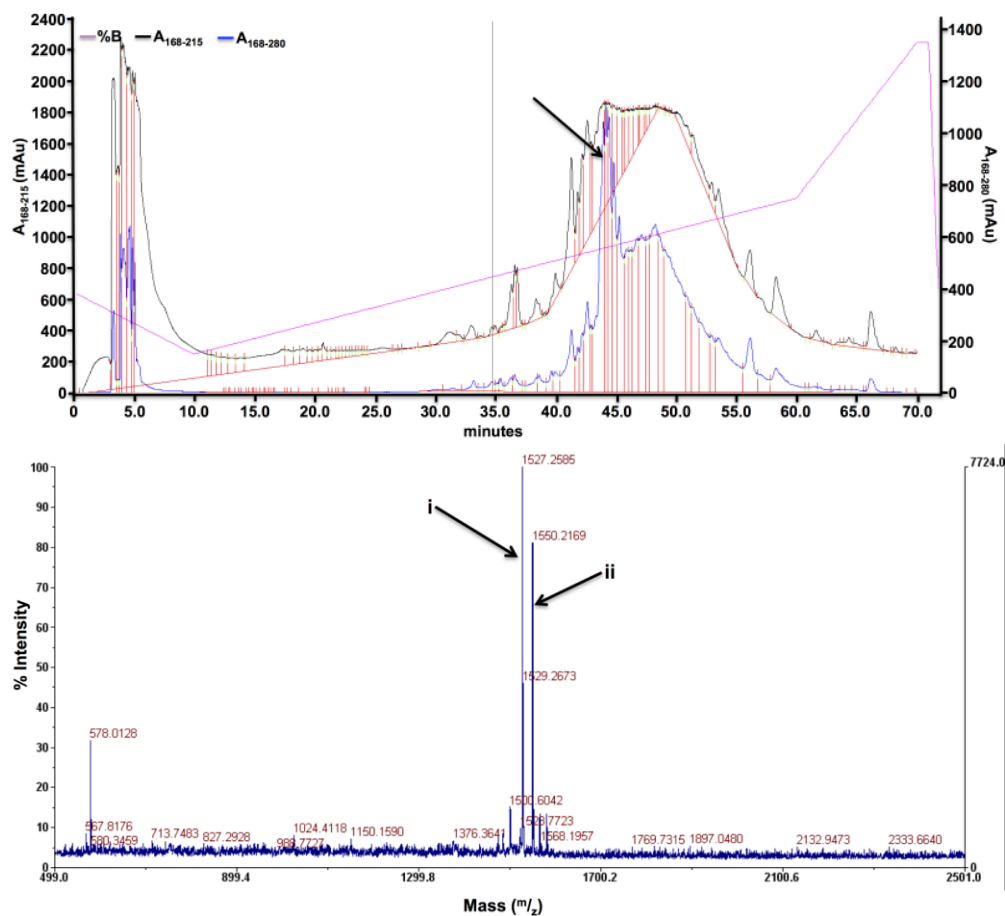




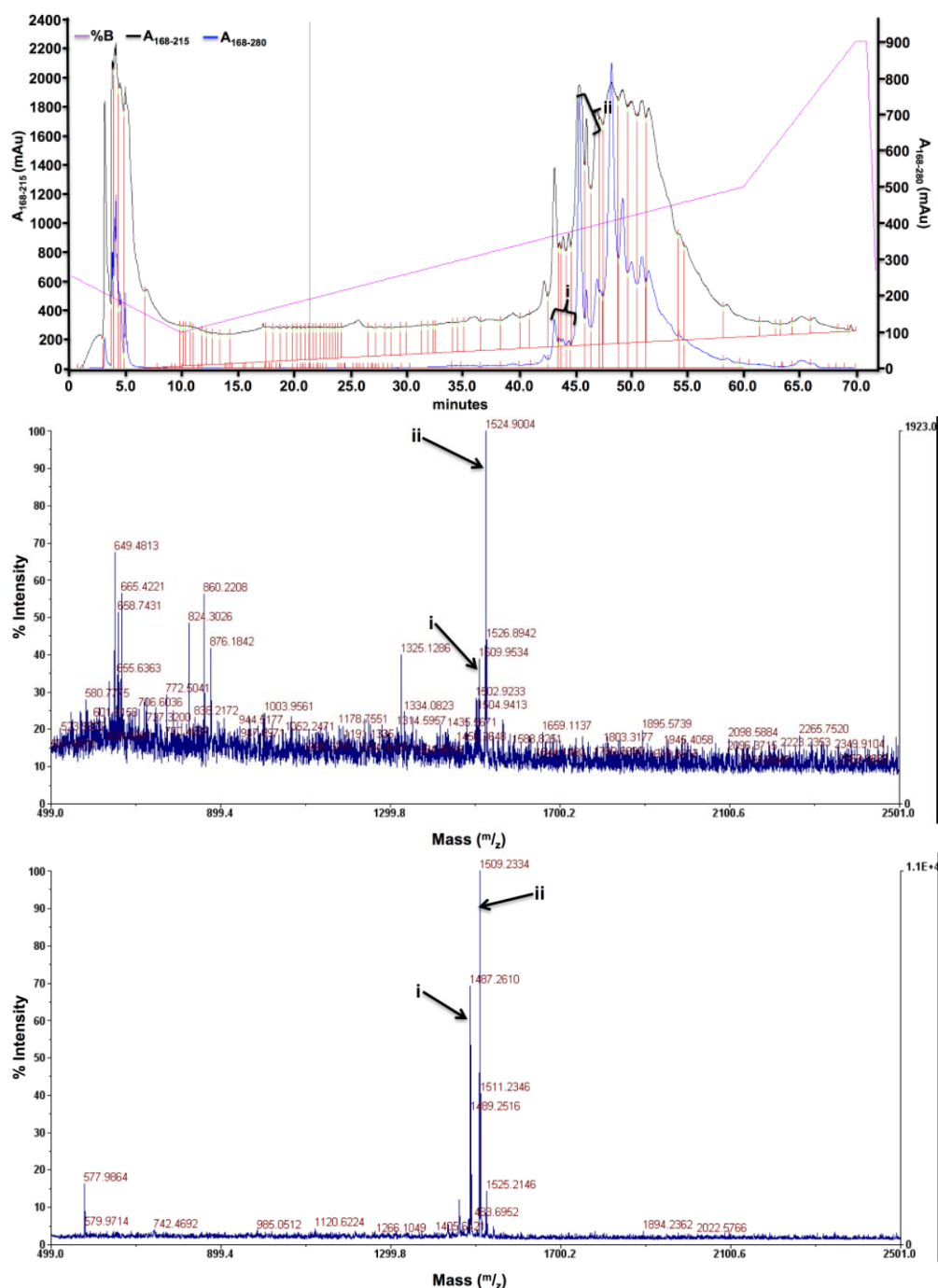
**Figure A3B.21. Characterization spectra for ligand hL2fF2.** Ligand **hL2fF2**:  $C_{72}H_{103}FN_{16}O_{18}S$ . (Top) The HPLC trace for ligand **hL2fF2**. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The peaks labeled “i” correspond to the  $[M+H]^+$  peak, expected  $[M+H] = 1531.742$ , and the peaks labeled “ii” correspond to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1553.724$ .



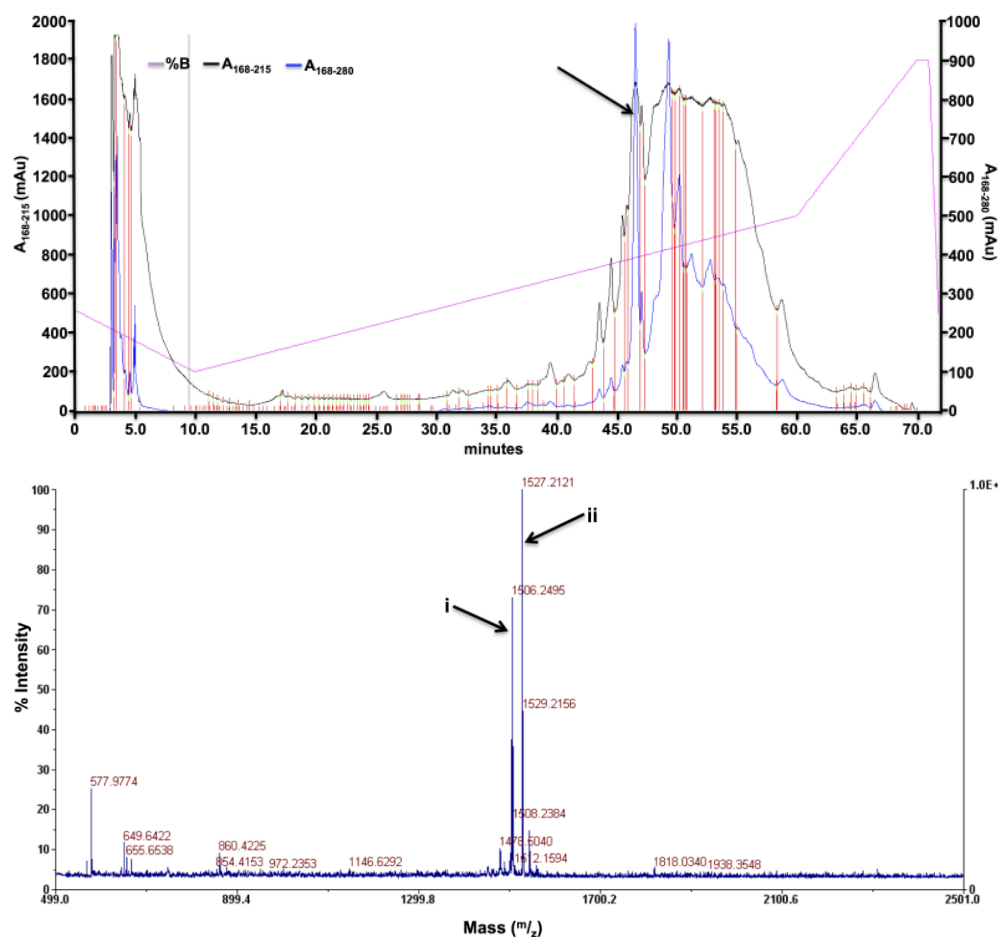
**Figure A3B.22. Representative characterization spectra for ligand hL2f\*F2.** Ligand hL2f\*F2:  $C_{72}H_{102}F_{21}N_{16}O_{18}S$ . (Top) The HPLC trace for ligand hL2f\*F2. The indicated regions correspond to the two fractions that had the correct mass. (Middle) The MALDI spectra for fraction 2. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H] = 1549.733$ , and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1571.714$ . (Bottom) The MALDI spectra for fraction 3. The indicated peak corresponds to the  $[M+Na]^+$  peak.



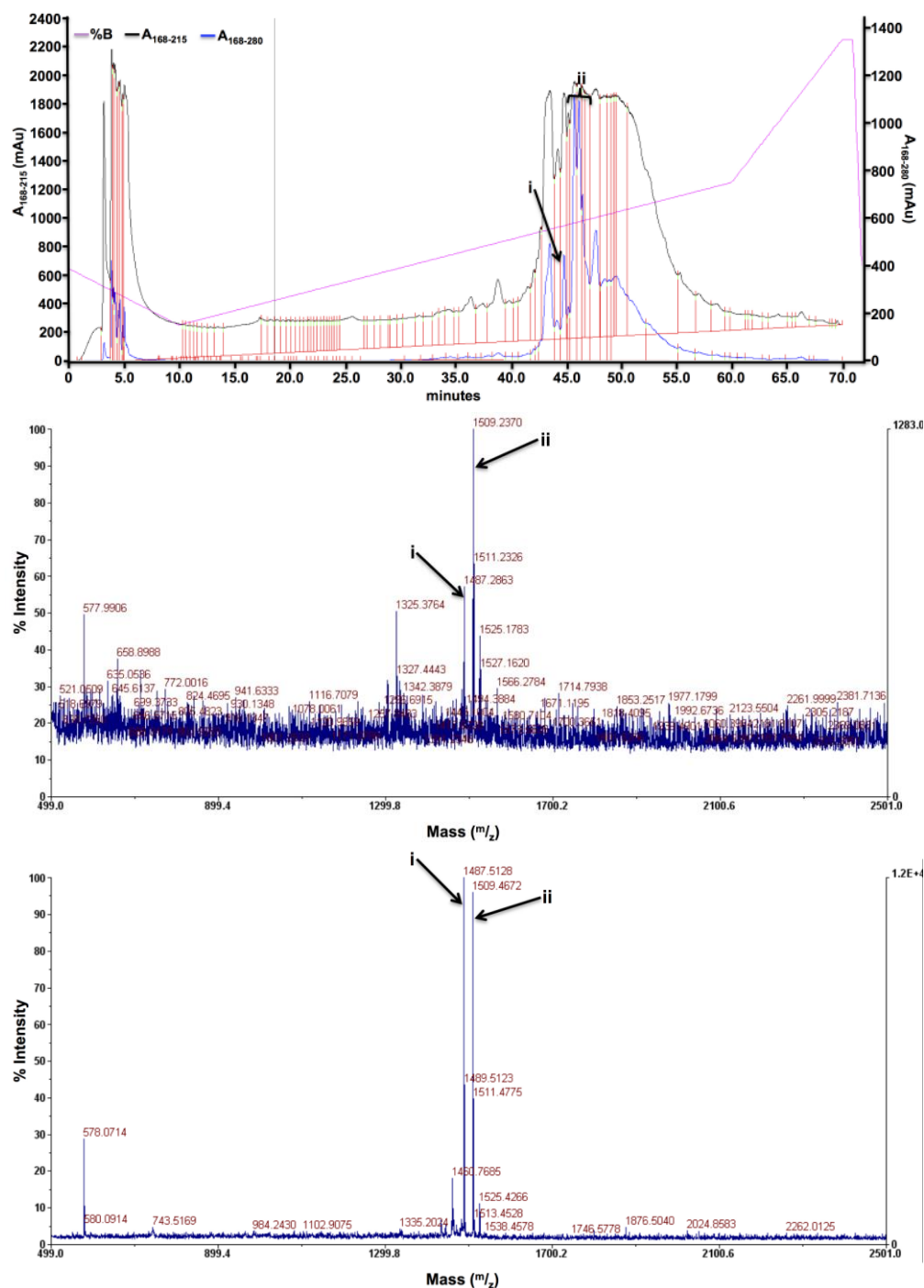
**Figure A3B.23. Characterization spectra for ligand hL3V2.** Ligand hL3V2:  $C_{74}H_{109}N_{15}O_{18}S$ . (Top) The HPLC trace for ligand hL3V2. The indicated regions correspond to the fraction that had the correct mass. (Bottom) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1528.787$ , and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na] = 1550.718$ .



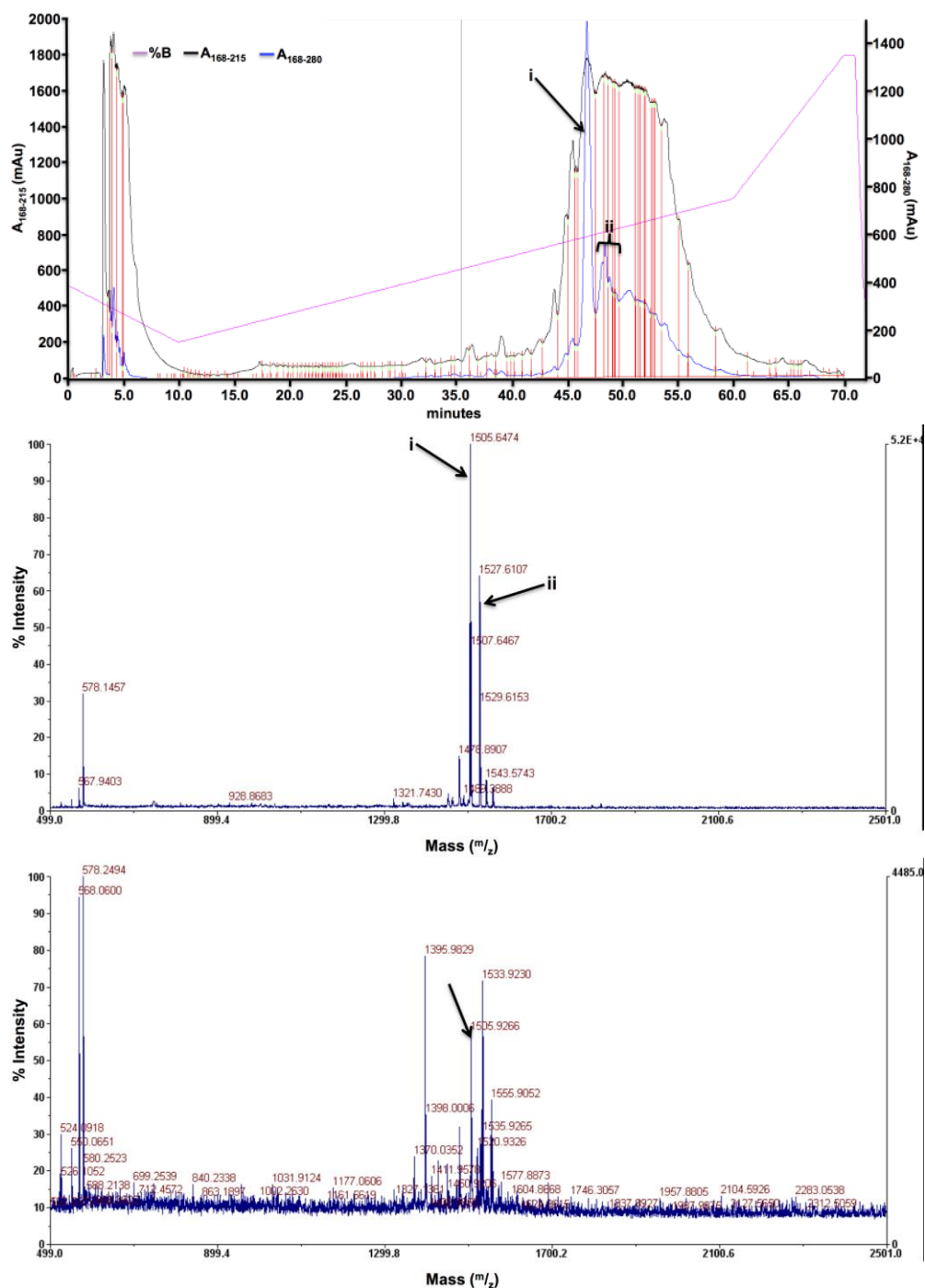
**Figure A3B.24. Characterization spectra for ligand **hL3fF3**.** Ligand **hL3fF3**:  $C_{71}H_{102}FN_{15}O_{17}S$ . (Top) The HPLC trace for ligand **hL3fF3**. The indicated regions correspond to the two fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+Na]$  peak, expected  $[M+Na] = 1510.718$ , and the peak labeled “ii” corresponds to the  $[M+K]$  peak, expected  $[M+K] = 1526.692$ . (Bottom) The MALDI spectra for fraction 2. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1488.736$ , and the peak labeled “ii” corresponds to the  $[M+Na]$  peak.



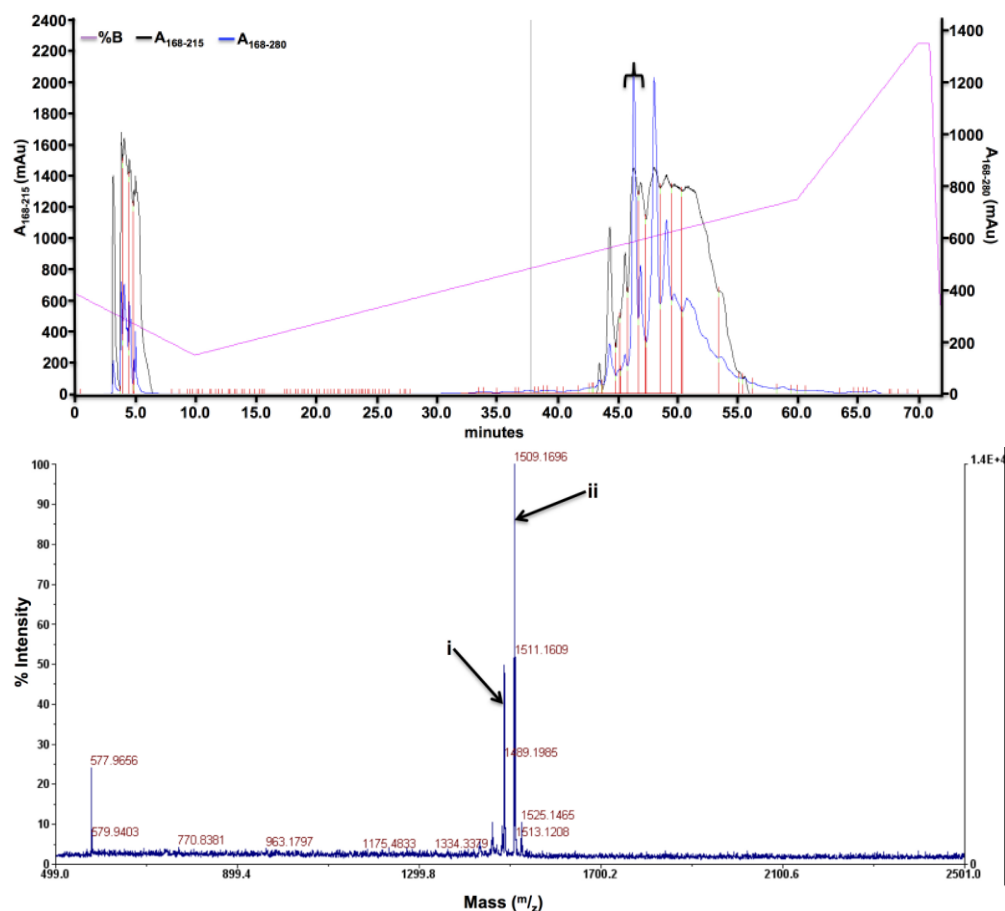
**Figure A3B.25. Characterization spectra for ligand hL3f\*F3.** Ligand hL3f\*F3:  $C_{71}H_{101}F_2N_{15}O_{17}S$ . (Top) The HPLC trace for ligand hL3f\*F3. The indicated regions correspond to the fraction that had the correct mass. (Bottom) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H]^+ = 1506.727$ , and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak, expected  $[M+Na]^+ = 1528.709$ .



**Figure A3B.26. Characterization spectra for ligand **hL3fF4**.** Ligand **hL3fF4**:  $C_{71}H_{102}FN_{15}O_{17}S$ . (Top) The HPLC trace for ligand **hL3fF4**. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The peaks labeled “i” correspond to the  $[M+H]^+$  peak, expected  $[M+H] = 1488.736$ , and the peaks labeled “ii” correspond to the  $[M+Na]^+$  peak, expected  $[M+Na] = 1510.718$ .

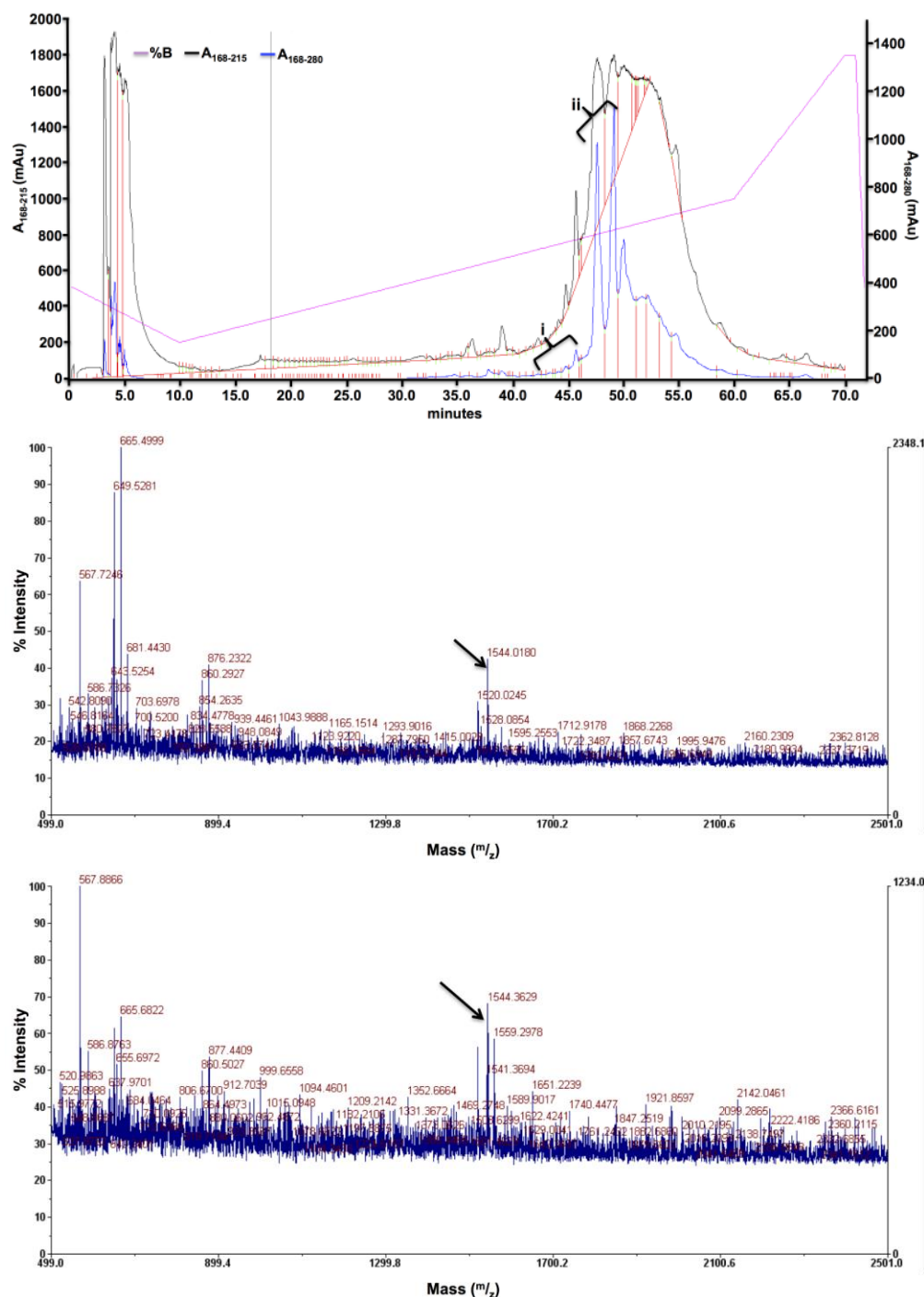


**Figure A3B.27. Characterization spectra for hL3f\*F4.** Ligand hL3f\*F4:  $C_{72}H_{102}F_2N_{16}O_{18}S$ . (Top) The HPLC trace for ligand hL3f\*F4. The indicated regions correspond to the two fractions that had the correct mass. (Middle) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1506.727$ , and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na] = 1528.709$ . (Bottom) The MALDI spectra for fraction 2. The indicated peak corresponds to the  $[M+Na]$  peak.

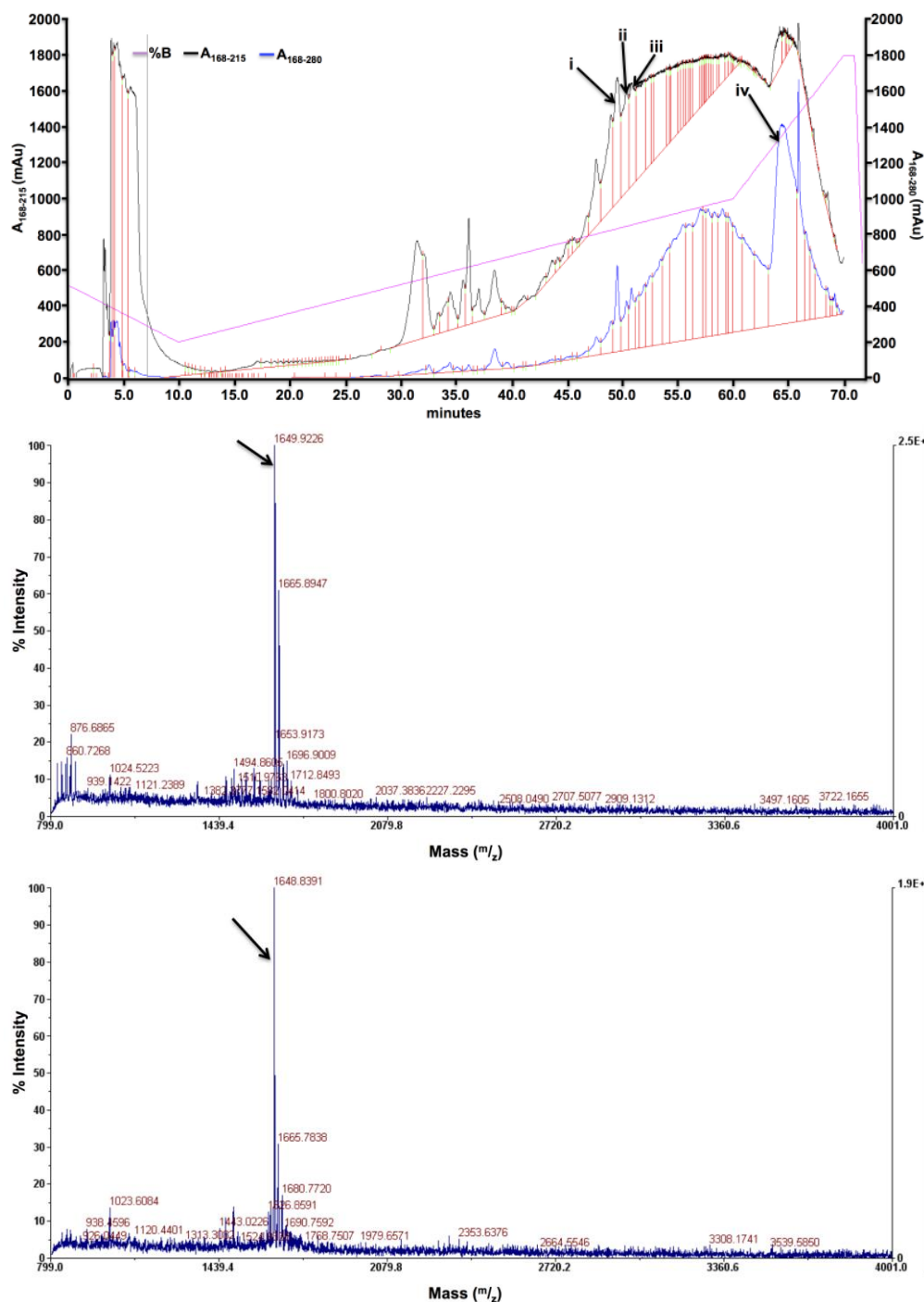


**Figure A3B.28. Characterization spectra for ligand hL3fF5.** Ligand hL3fF5:  $C_{71}H_{102}FN_{15}O_{17}S$ . (Top) The HPLC trace for ligand hL3fF5. The indicated regions correspond to the fraction that had the correct mass. (Bottom) The MALDI spectra for fraction 1. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1488.736$ , and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na] = 1510.718$ .

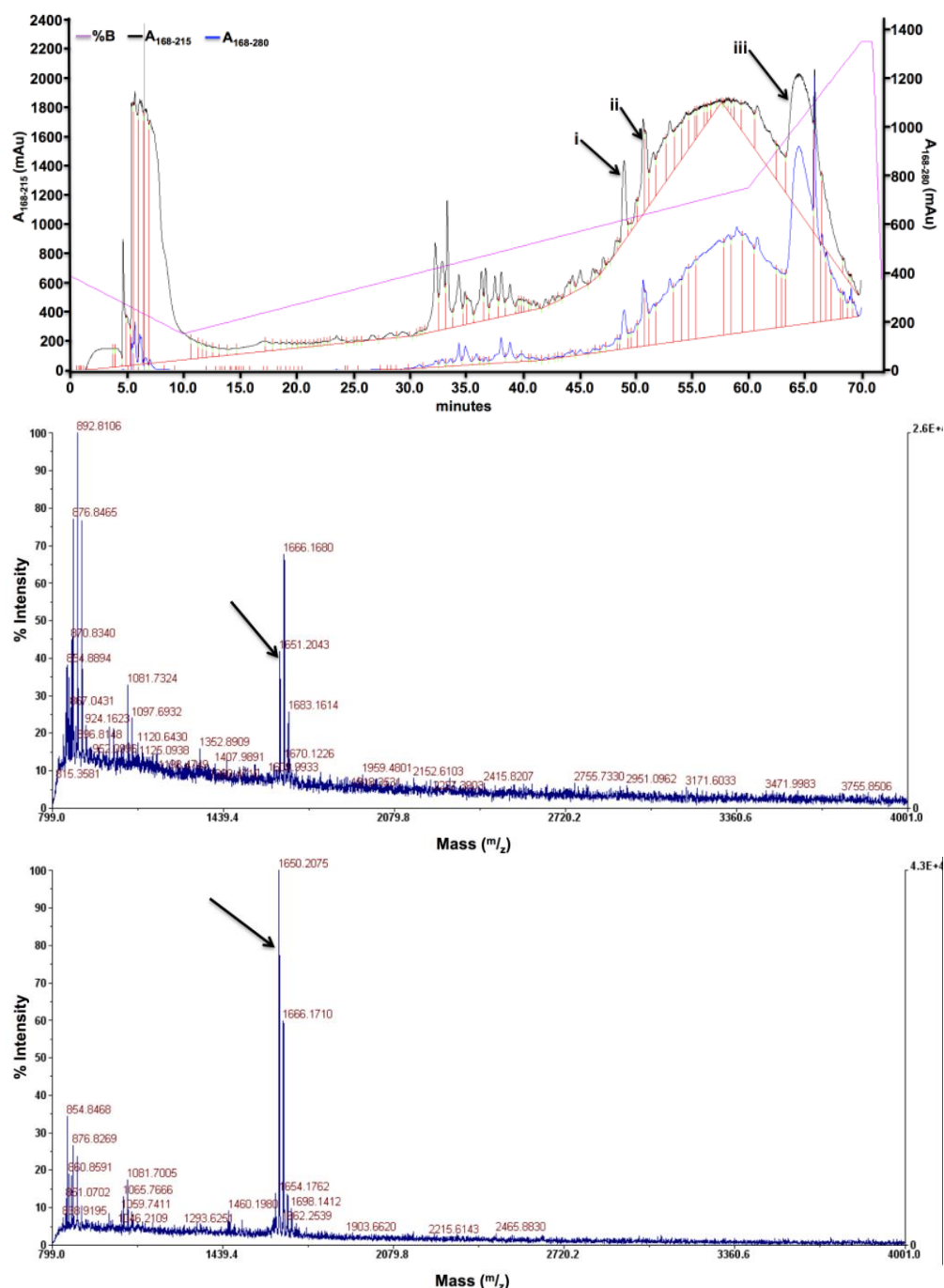




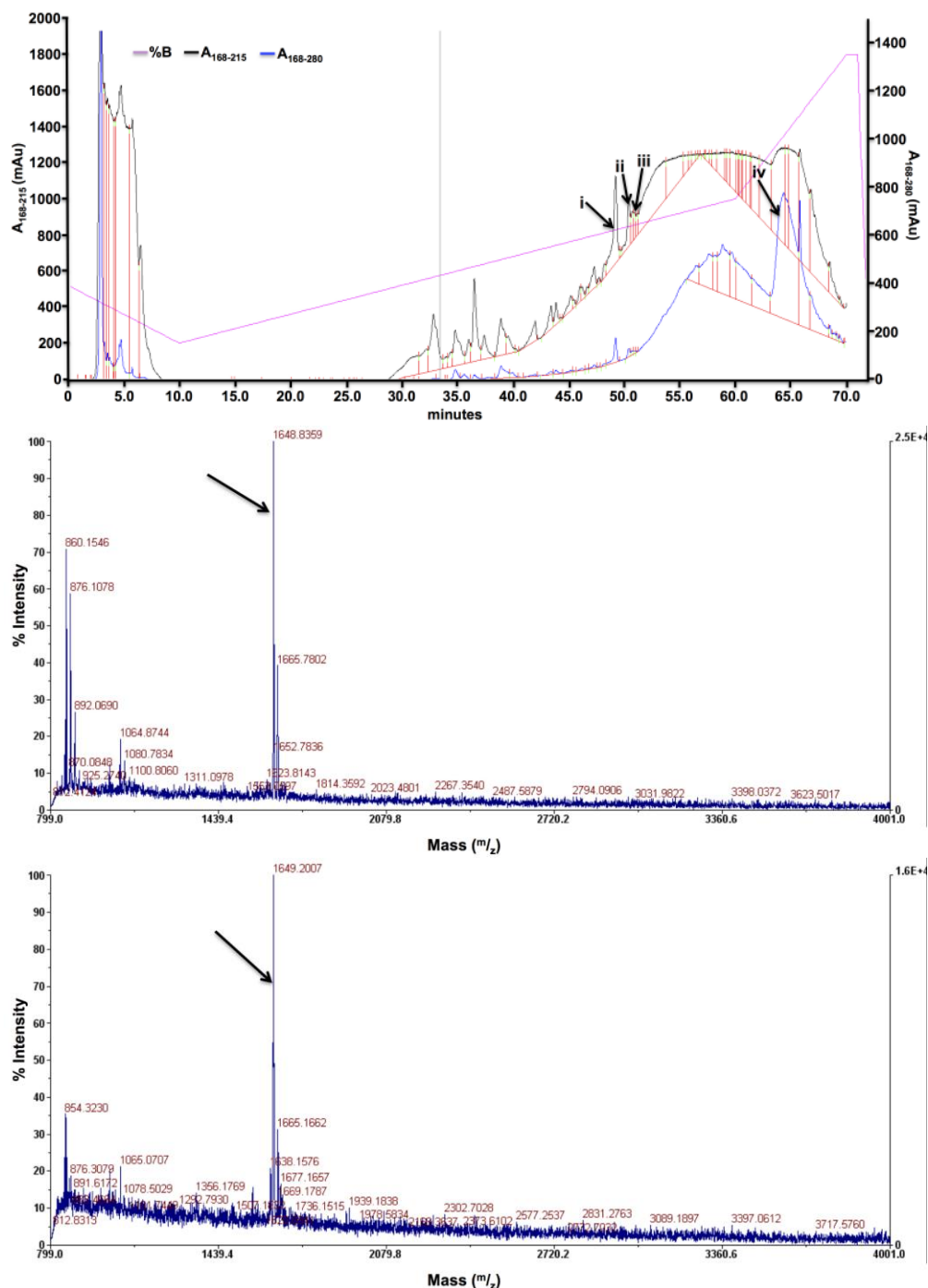
**Figure A3B.29. Characterization spectra for ligand hL3f\*F5.** Ligand hL3f\*F5:  $C_{71}H_{101}F_2N_{15}O_{17}S$ . (Top) The HPLC trace for ligand hL3f\*F5. The indicated regions correspond to the two fractions that had the correct mass. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The indicated peaks correspond to the  $[M+K]$  peak, expected  $[M+K] = 1544.683$ .



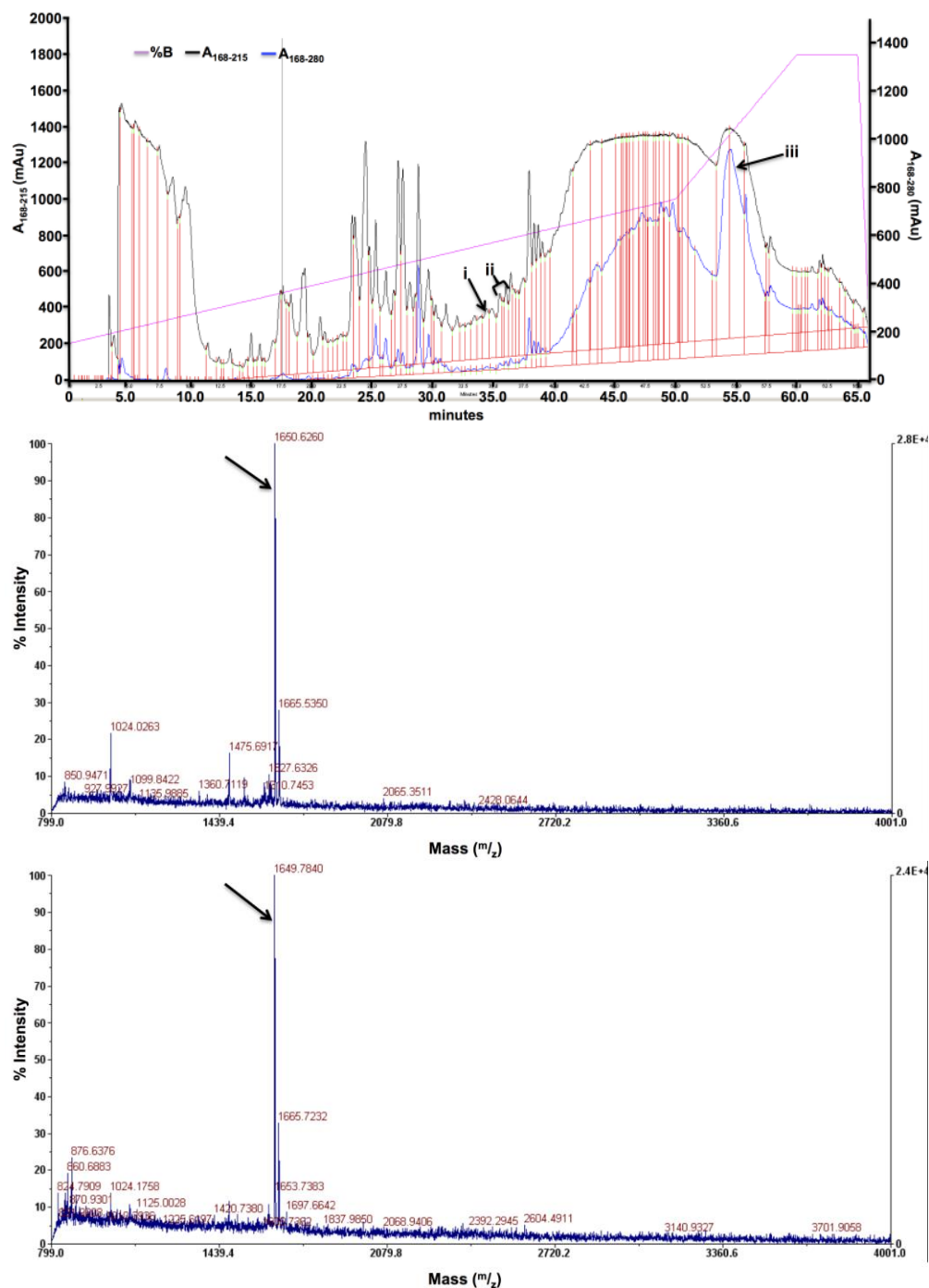
**Figure A3B.30. Representative characterization spectra for ligand Ak-hL3G1V3.** Ligand **Ak-hL3G1V3**:  $C_{82}H_{125}N_{13}O_{19}S$ . (Top) The HPLC trace for ligand **Ak-hL3G1V3**. The indicated peaks i-iii correspond to the three fractions that had the correct mass. The peak labeled “iv” corresponds to oxidized Grubbs I catalyst. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The indicated peaks correspond to the  $[M+Na]$  peak, expected  $[M+Na] = 1650.883$ .



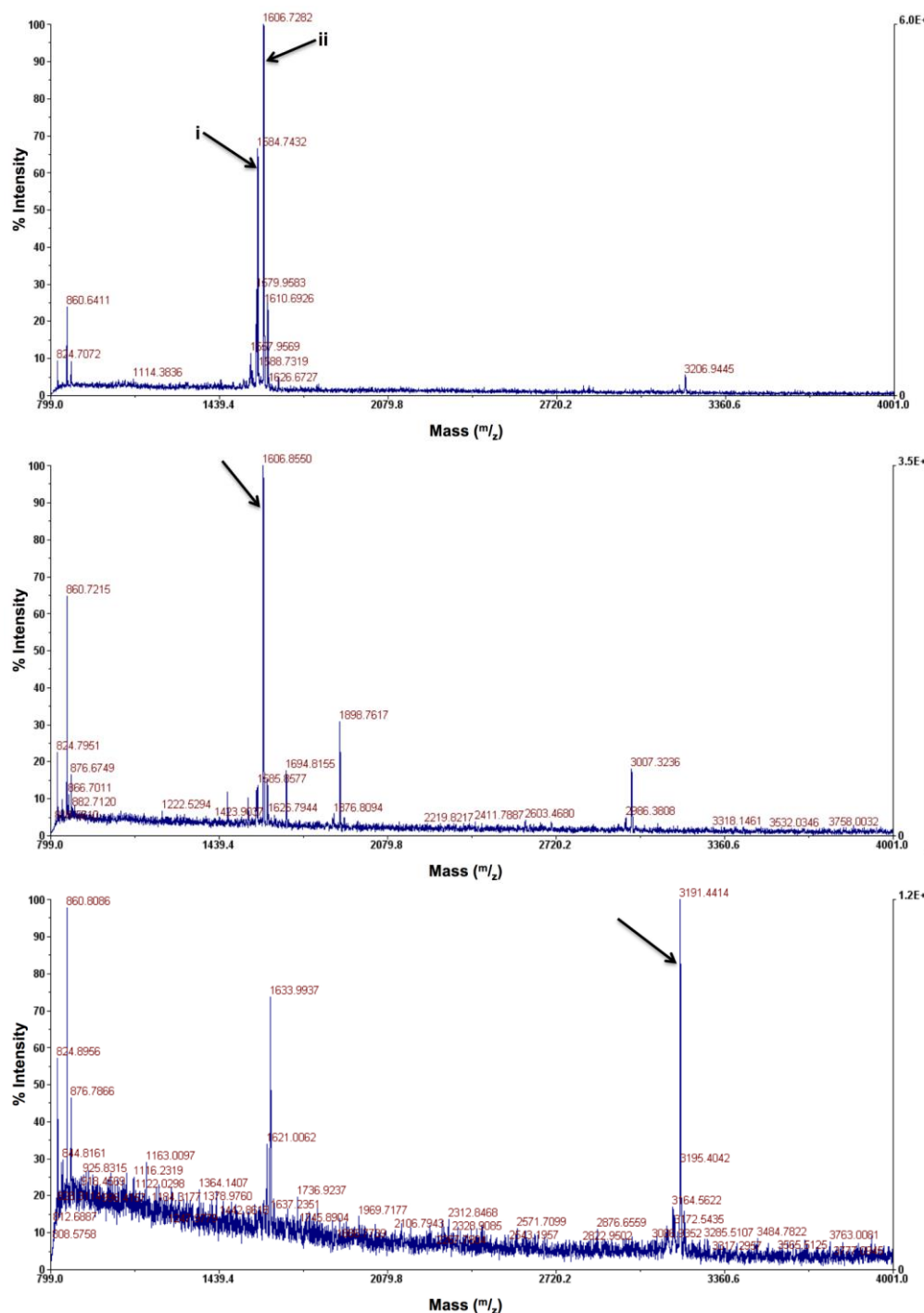
**Figure A3B.31. Characterization spectra for ligand Ak-hL3V2G6.** Ligand Ak-hL3V2G6:  $C_{82}H_{125}N_{13}O_{19}S$ . (Top) The HPLC trace for ligand Ak-hL3V2G6. The indicated peaks i-ii correspond to the two fractions that had the correct mass. The peak labeled “iii” corresponds to oxidized Grubbs I catalyst. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The indicated peaks correspond to the  $[M+Na]$  peak, expected  $[M+Na] = 1650.883$ .



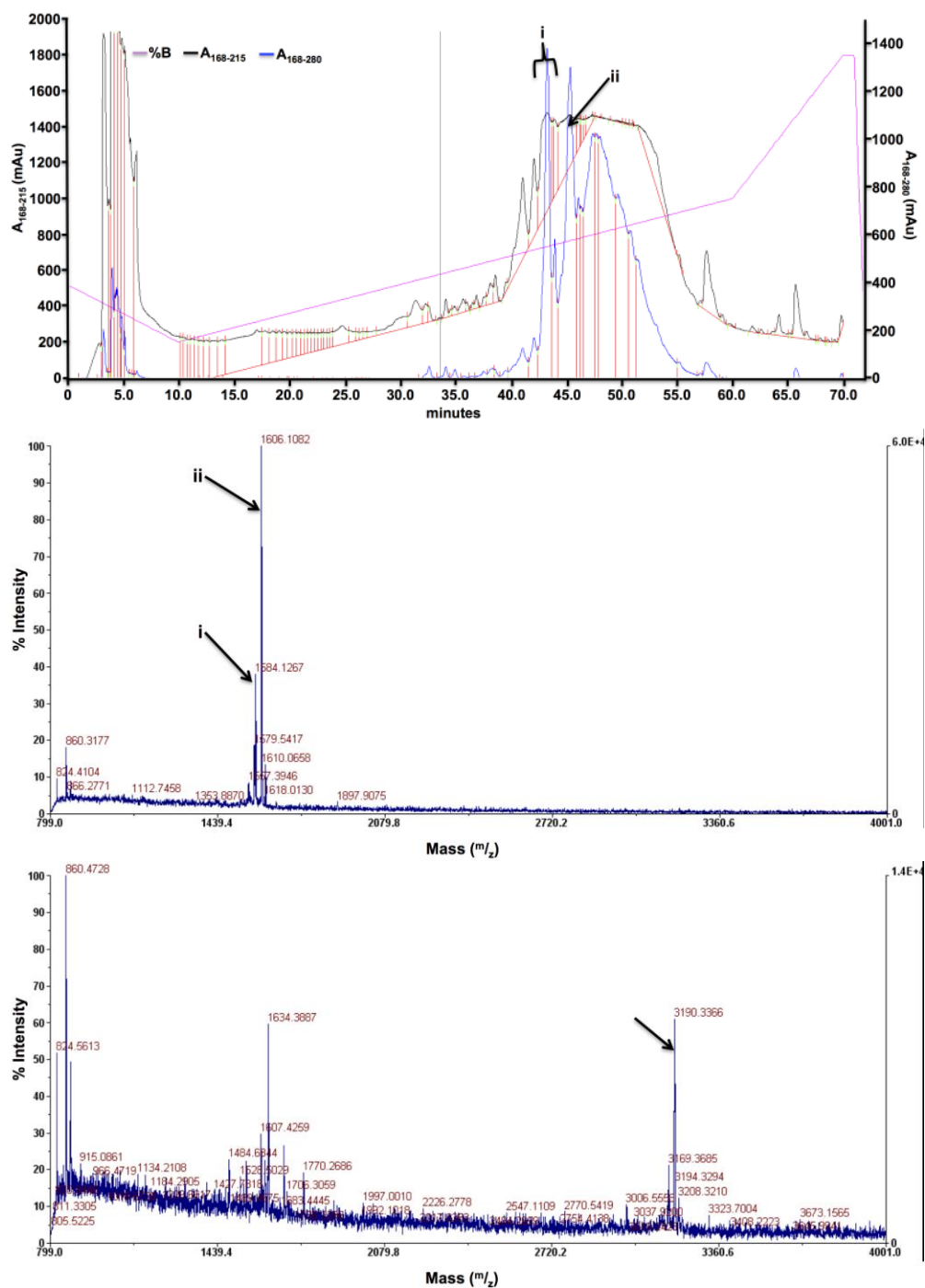
**Figure A3B.32. Representative characterization spectra for ligand Ak'-hL3G1V3.** Ligand Ak'-hL3G1V3:  $C_{82}H_{125}N_{13}O_{19}S$ . (Top) The HPLC trace for ligand Ak'-hL3G1V3. The indicated peaks i-iii correspond to the three fractions that had the correct mass. The peak labeled "iv" corresponds to oxidized Grubbs I catalyst. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The indicated peaks correspond to the  $[M+Na]$  peak, expected  $[M+Na] = 1650.883$ .



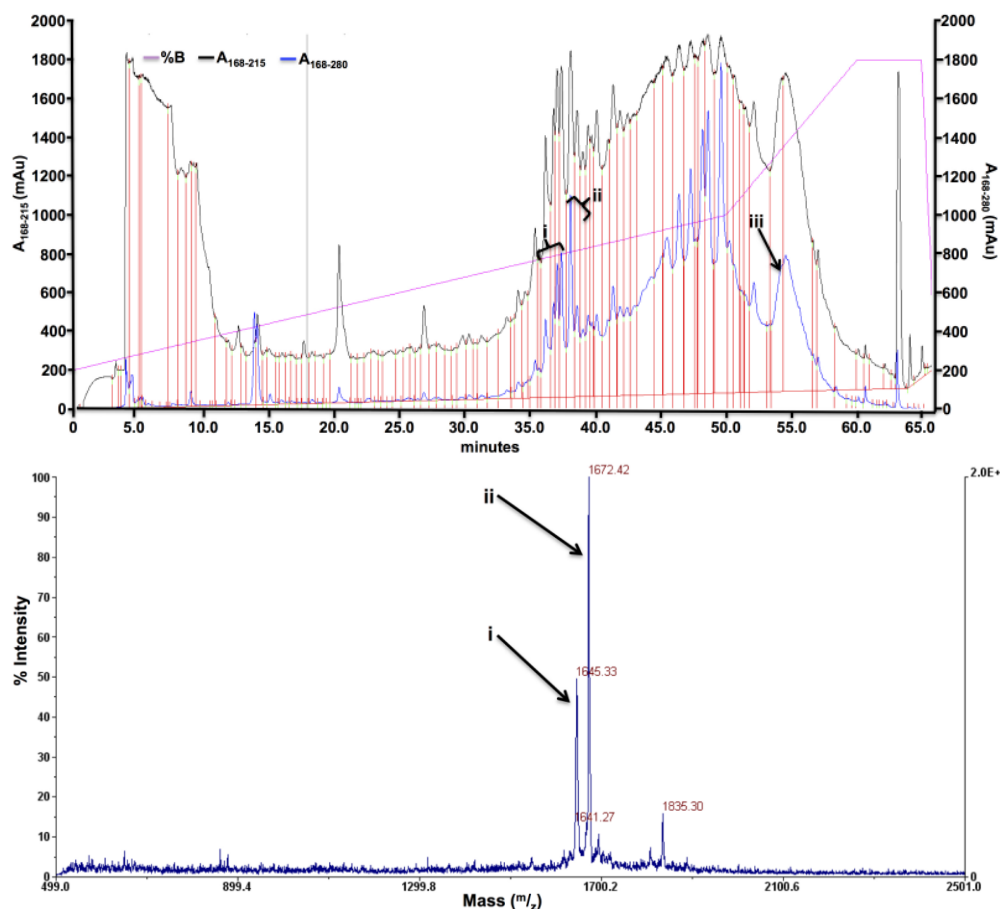
**Figure A3B.33. Characterization spectra for ligand Ak'-hL3V2G6.** Ligand Ak'-hL3V2G6:  $C_{82}H_{125}N_{13}O_{19}S$ . (Top) The HPLC trace for ligand Ak'-hL3V2G6. The indicated peaks i-ii correspond to the two fractions that had the correct mass. The peak labeled "iii" corresponds to oxidized Grubbs I catalyst. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The indicated peaks correspond to the  $[M+Na]$  peak, expected  $[M+Na] = 1650.883$ .



**Figure A3B.34. MALDI spectra for ligand **Tz-hL3G1V3**.** Ligand **Tz-hL3G1V3**:  $C_{76}H_{112}N_{16}O_{19}S$  and  $C_{152}H_{224}N_{32}O_{38}S_2$  (dimer). (Top) The MALDI trace for fraction 1, which is mostly monomer. The peak labeled “i” corresponds to the  $[M+H]$  peak, expected  $[M+H] = 1585.809$ , and the peak labeled “ii” corresponds to the  $[M+Na]$  peak, expected  $[M+Na] = 1607.791$ . (Middle) The MALDI spectra for fraction 2, which is mostly monomer. The indicated peak corresponds to the  $[M+Na]$  peak. (Bottom) The MALDI spectra for fraction 3, which is the dimeric form of ligand **Tz-hL3G1V3**. The indicated peaks correspond to the  $[D+Na]$  peak, expected  $[D+Na] = 3191.441$ .

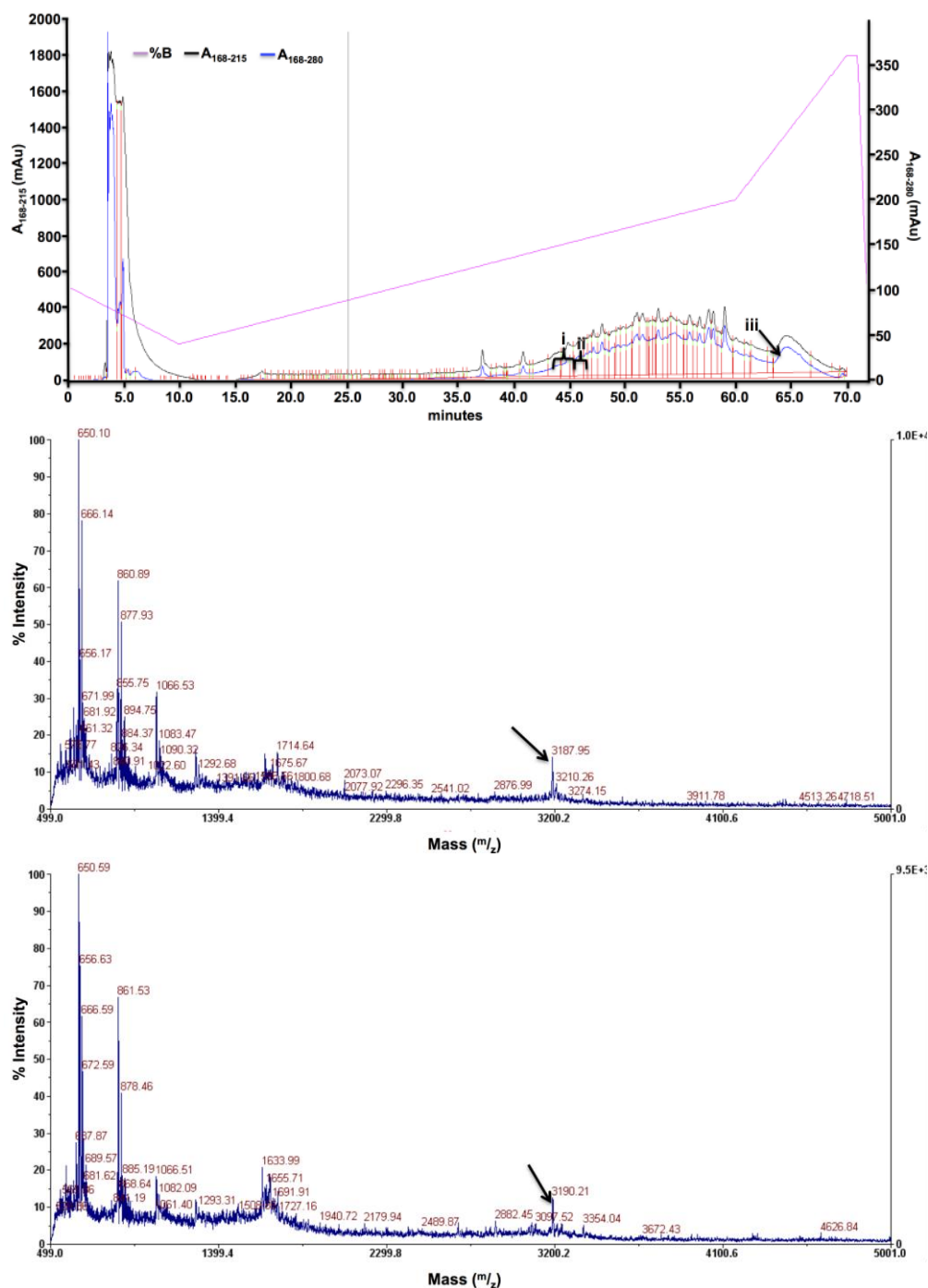


**Figure A3B.35. Characterization spectra for ligand **Tz-hL3V2G6**.** Ligand **Tz-hL3V2G6**:  $C_{76}H_{112}N_{16}O_{19}S$  and  $C_{152}H_{224}N_{32}O_{38}S_2$  (dimer). (Top) The HPLC trace for ligand **Tz-hL3V2G6**. The indicated peaks correspond to the two fractions with the correct mass. (Middle) The MALDI trace for fraction 1, which is all monomer. The peak labeled “i” corresponds to the  $[M+H]^+$  peak, expected  $[M+H]^+ = 1585.809$ , and the peak labeled “ii” corresponds to the  $[M+Na]^+$  peak, expected  $[M+Na]^+ = 1607.791$ . (Bottom) The MALDI spectra for fraction 3, which is the dimeric form of ligand **Tz-hL3V2G6**. The indicated peak corresponds to the  $[D+Na]^+$  peak, expected  $[D+Na]^+ = 3192.592$ .



**Figure A3B.36. Representative characterization spectra for C-term ligand.** Ligand **C-term**:  $C_{76}H_{120}N_{24}O_{17}S$ . (Top) The HPLC trace for ligand **C-term**. The indicated regions i-ii correspond to the two fractions that had the correct mass. The peak labeled “iii” corresponds to oxidized Grubbs I catalyst. (Bottom) The MALDI spectra for fraction 2. The peak labeled “i” corresponds to the  $[M-N_2+H]$  peak, expected  $[M-N_2+H] = 1647.916$ , and the peak labeled “ii” corresponds to the  $[M+H]$  mass, expected  $[M+H] = 1673.906$ .





**Figure A3B.37. Characterization spectra for ligand DNBCtermH8.** Ligand **DNBCtermH8**:  $C_{150}H_{226}N_{44}O_{32}S$ . (Top) The HPLC trace for ligand **DNBCtermH8**. The indicated peaks i-ii correspond to the two fractions that had the correct mass. The peak labeled “iii” corresponds to oxidized Grubbs I catalyst. (Middle, Bottom) The MALDI spectra for fraction 1 and fraction 2 respectively. The indicated peaks correspond to the  $[M+H]$  peak, expected  $[M+H] = 3188.721$ .



## Appendix 3C

### Supplemental Tables

#### Table of Contents

Table A3C.1-Alanine scan sets.	166
Table A3C.2-Generation two ligands.	166
Table A3C.3-Alternative variants of <b>hL3V3</b> prepared.	166
Table A3C.4-Linker hits from JXL1 vs. N-term linker screen.	166
Table A3C.5-Linker hits from DNB vs. C-term linker screen.	167
Table A3C.6-Linker hits from the DNB vs. N-term linker screen.	167
Table A3C.7-Linker hits from JXL1 vs. DNB linker screen.	167
Table A3C.8-Linker hits from JXL1 vs. C-term linker screen.	167
Table A3C.9-Ligand characterization data.	168
References	174

Table A3C.1. The Alanine scan sets for the ligands.

hL1	hL2	hL3
AYQLL (hL1a1)	AYYNV (hL2a1)	AGYYY (hL3a1)
YAQLL (hL1a2)	YAYNV (hL2a2)	LAYYY (hL3a2)
YYALL (hL1a3)	YYANV (hL2a3)	LGAYY (hL3a3)
YYQAL (hL1a4)	YYYAV (hL2a4)	LGYAY (hL3a4)
YYQLA (hL1a5)	YYYNA (hL2a5)	LGYYA (hL3a5)

Table A3C.2. The generation two ligands.

hL1	hL2	hL3	
YYELL (hL1E3)	<sup>f</sup> FYYNV (hL2 <sup>f</sup> F1)	LVYYY (hL3V2)	LG <sup>3,4-f</sup> FYY (hL3 <sup>f*</sup> F3)
YYDLL (hL1D3)	<sup>f</sup> FYYNV (hL2 <sup>f</sup> F2)	LG <sup>f</sup> FYY (hL3 <sup>f</sup> F3)	LG <sup>Y<sup>3,4-f</sup></sup> FY (hL3 <sup>f*</sup> F4)
	<sup>3,4-f</sup> FYYNV (hL2 <sup>f*</sup> F1)	LG <sup>f</sup> FY (hL3 <sup>f</sup> F3)	LGYY <sup>3,4-f</sup> F (hL3 <sup>f*</sup> F5)
	<sup>Y<sup>3,4-f</sup></sup> FYYNV (hL2 <sup>f*</sup> F2)	LGYY <sup>f</sup> F (hL3 <sup>f</sup> F5)	

Table A3C.3. Alternative variants of hL3V2 prepared.

Triazole Closure	RCM Closure	
GLVYYY (Tz-hL3G1V3)	OctGLVYYYPen (Ak-hL3G1V3)	PenGLVYYYOct (Ak'-hL3G1V3)
LVYYYG (Tz-hL3V2G6)	OctLVYYYGPen (Ak-hL3V2G6)	PenLVYYYGOct (Ak'-hL3V2G6)

Table A3C.4. Linker hits from JXL1 vs. N-term linker screen.

Hit	Click	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	L-Pra	p	l	G	G	
2	L-Pra	p	l	Aib	Aib	
3	L-Pra	Aib	l	G	G	
4	-	l	Aib	Aib	G	l
5	L-Pra	p	l	G	p	l
6	L-Pra	l	p	l		
7	L-Pra	l	l	Aib		

**Table A3C.5. Linker hits from DNB vs. C-term linker screen.**

Hit	Click	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	L-Pra	G	l	G	l	G
2	L-Pra	p	l	Aib	Aib	
3	L-Pra	l	p	l	G	
4	L-Pra	l	G	l	p	
5	L-Pra	p	Aib	G	G	
6	L-Pra	p	l	G		
7	L-Pra	l	l	Aib	Aib	
8	L-Pra	p	G	l	l	

**Table A3C.6. Linker hits from the DNB vs. N-term linker screen.**

Hit	Click	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	L-Pra	l	G	l	G	l
2	L-Pra	G	Aib	Aib	G	
3	-	l	l	G	l	
4	L-Pra	p	Aib	G	Aib	
5	L-Pra	p	l	Aib		
6	L-Pra	p	G	p	Aib	Aib
7	L-Pra	l	G	G		

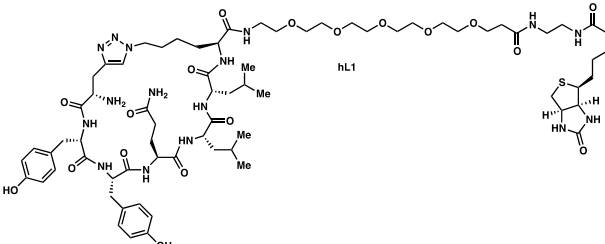
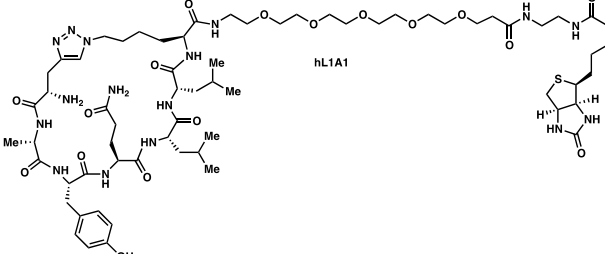
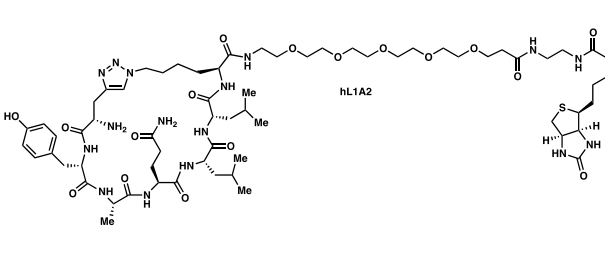
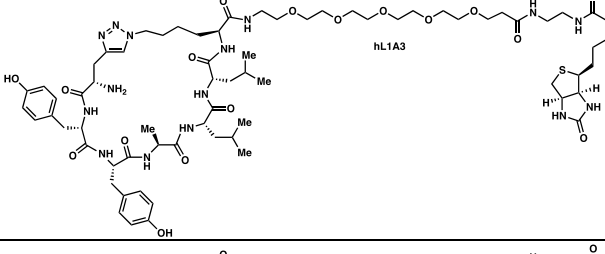
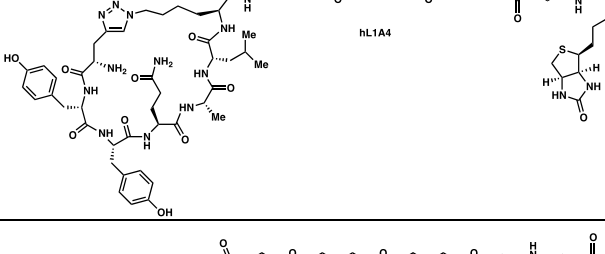
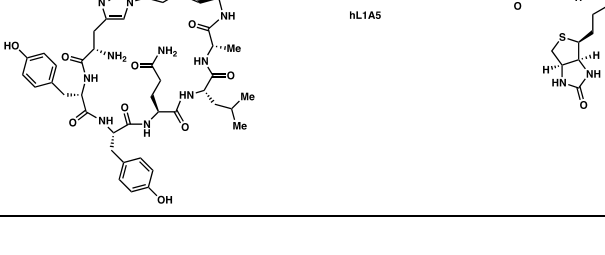
**Table A3C.7. Linker hits from JXL1 vs. DNB linker screen.**

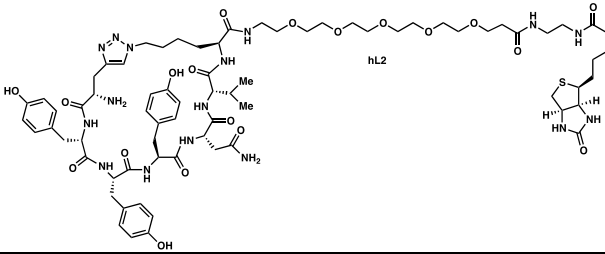
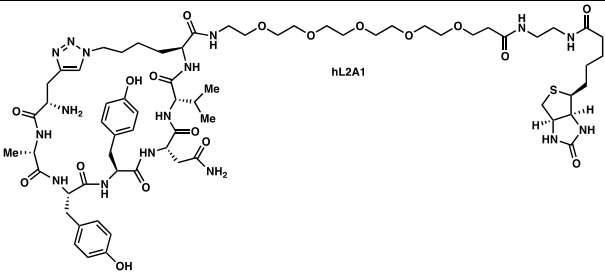
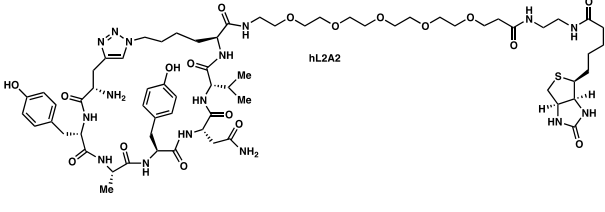
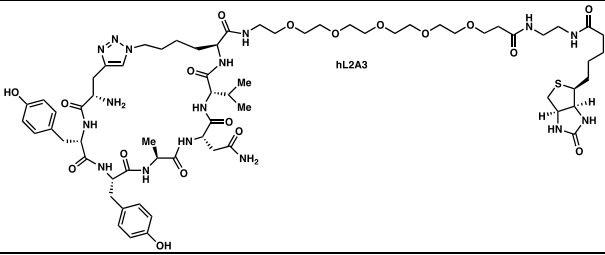
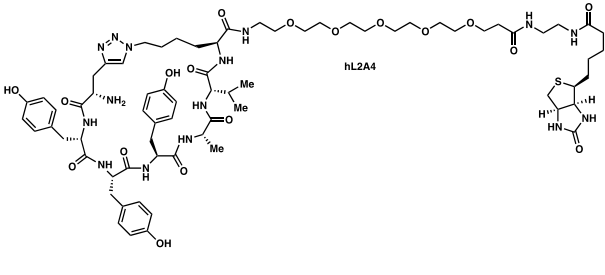
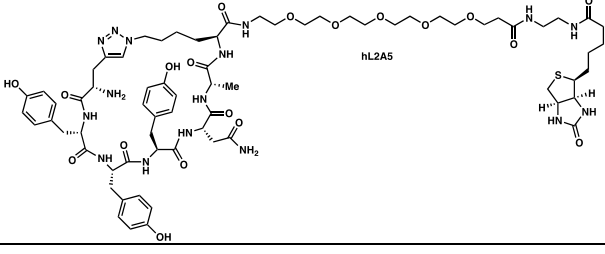
Hit	Click	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	L-Pra	p	l			
2	L-Pra	G	p	Aib	Aib	
3	L-Pra	l	p	l	Aib	Aib

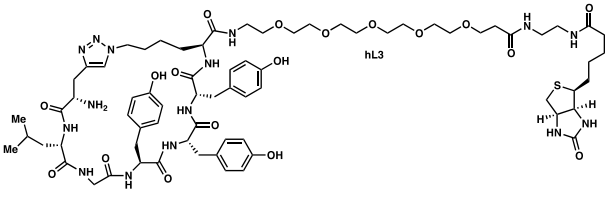
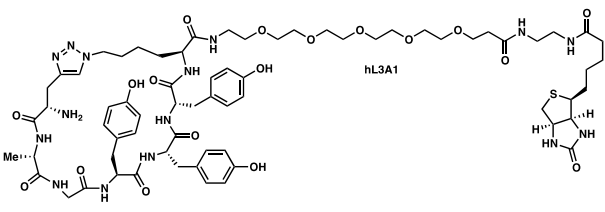
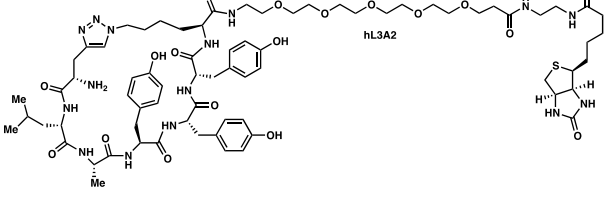
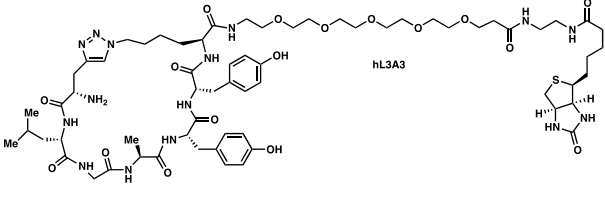
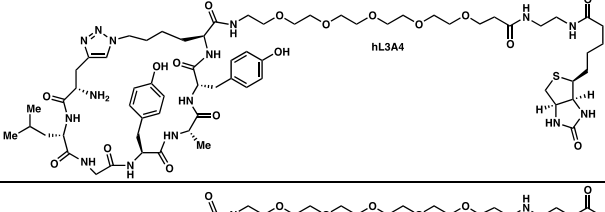
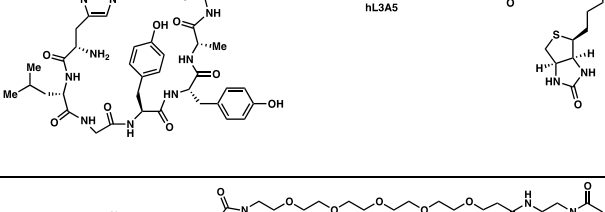
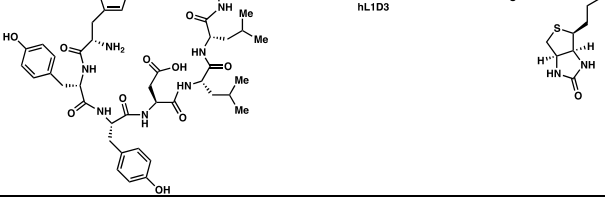
**Table A3C.8. Linker hits from JXL1 vs. C-term linker screen.**

Hit	Click	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>
1	L-Pra	Aib	G	G	l	Aib
2	L-Pra	l	G	Aib	G	l
3	L-Pra	l	l			l
4	L-Pra	G	p	l	l	G

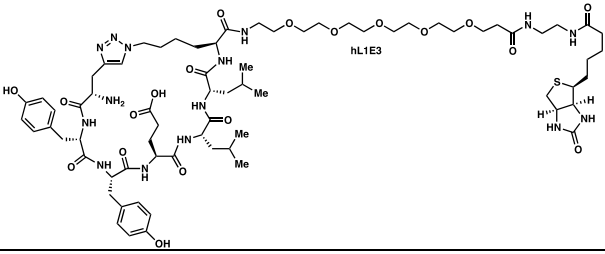
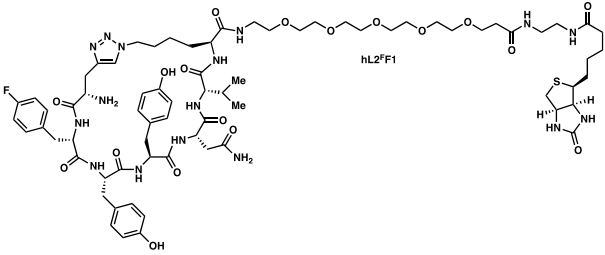
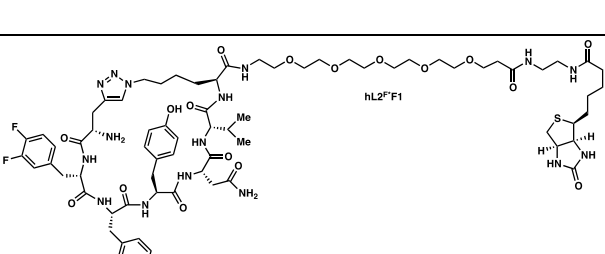
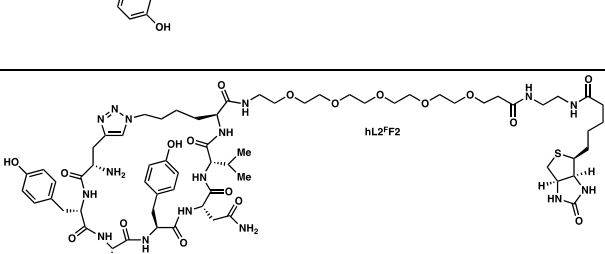
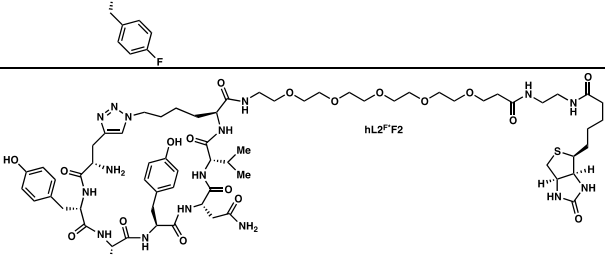
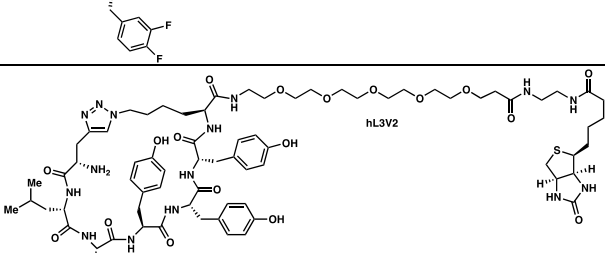
Table A3C.9. Ligand Characterization Data.<sup>1</sup>

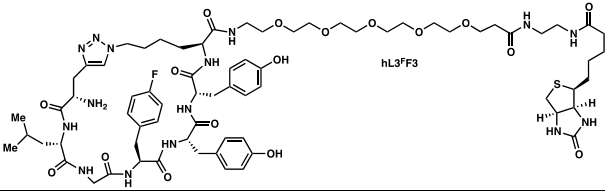
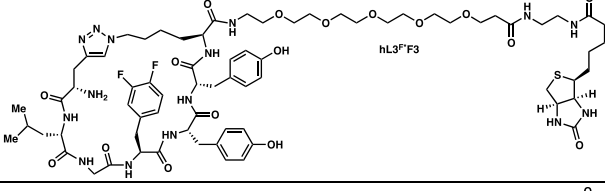
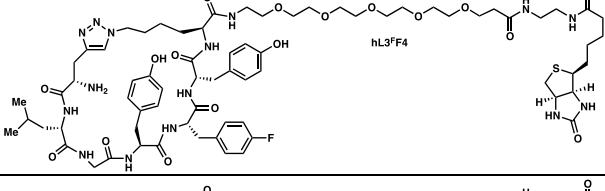
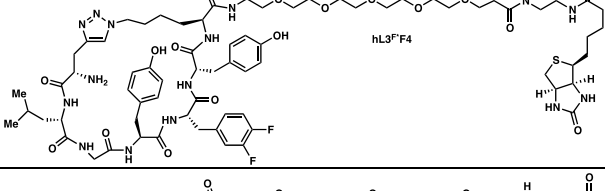
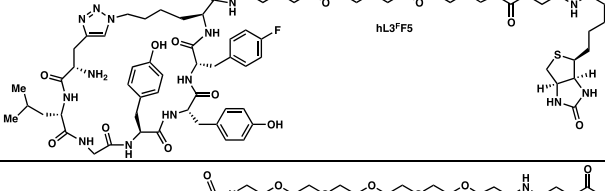
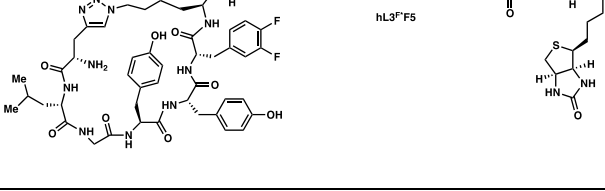
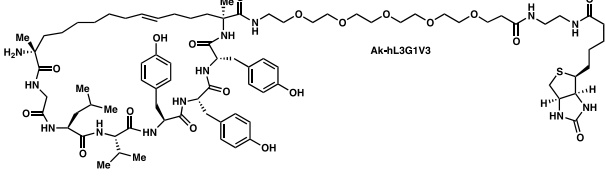
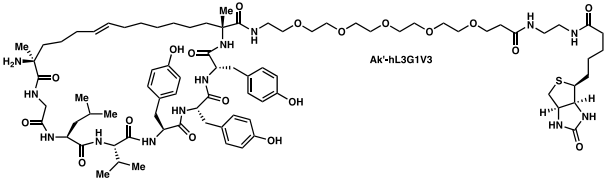
Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
	$C_{71}H_{110}N_{16}O_{18}S$	*	[M+H]= 1506.790	*
	$C_{65}H_{106}N_{16}O_{17}S$	46.5-47 (f1); 47-48.5 (f2)	[M+H]= 1416.677; [M+Na]= 1438.762	1414.613; 1436.605 (f1) 1436.962 (f2)
	$C_{65}H_{106}N_{16}O_{17}S$	32.5- 33.5 (f1); 33.5- 34.5 (f2); 34.5- 35.5 (f3); 35.5-36 (f4)	[M+H]= 1416.677; [M+Na]= 1438.762	1415.510; 1437.511 (f1) 1415.678; 1437.650 (f2) 1415.836; 1437.838 (f3) 1415.975; 1437.974 (f4)
	$C_{69}H_{107}N_{15}O_{17}S$	46.5-48 (f1); 48-48.5 (f2); 48.5-50 (f3)	[M+H]= 1450.777; [M+Na]= 1472.759	1450.047 (f1) 1450.283; 1472.253 (f2) 1450.467; 1472.442 (f3)
	$C_{68}H_{104}N_{16}O_{18}S$	41:30- 44 (f1); 44-46.5 (f2)	[M+H]= 1465.751; [M+Na]= 1487.733; [M+K]= 1503.707	1464.937; 1502.864 (f1) 1465.173; 1487.151 (f2)
	$C_{68}H_{104}N_{16}O_{18}S$	40-43 (f1); 43-45 (f2); 45-46.5 (f3); 46.5-49 (f4)	[M+H]= 1465.751; [M+Na]= 1487.733; [M+K]= 1503.707	1465.417; 1487.401; 1503.373 (f1) 1488.548 (f2) 1465.313; (f3) 1466.837; 1487.835 (f4)

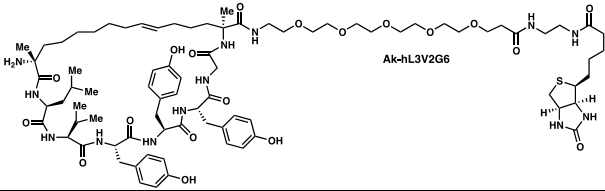
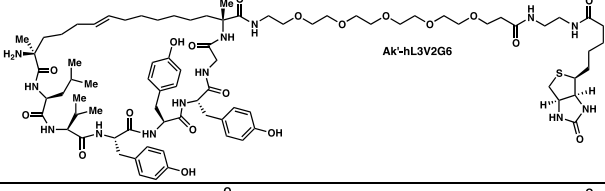
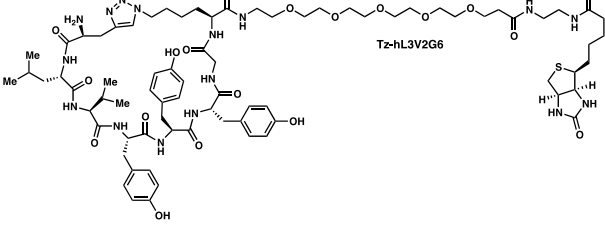
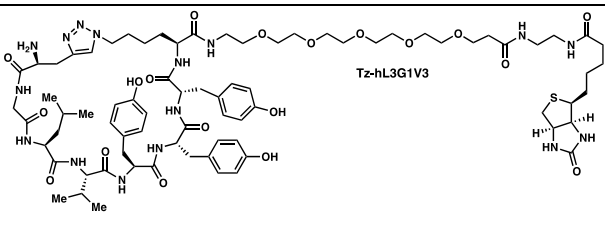
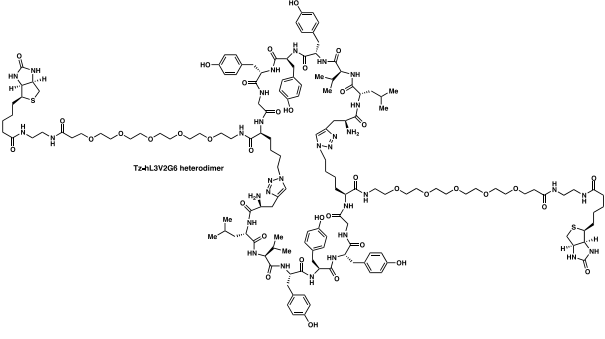
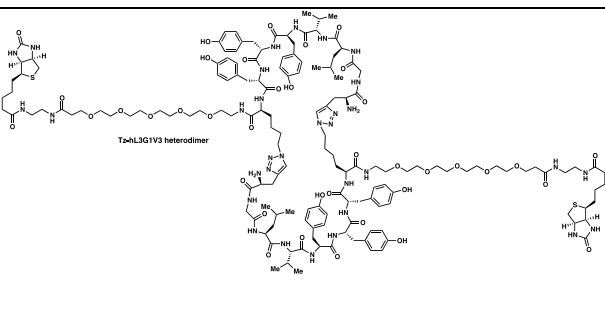
Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
 hL2	$C_{72}H_{104}N_{16}O_{19}S$	37-40 (f1); 40-44 (f2)	[M+H]= 1530.765	1529.177 (f1) 1529.348; 1551.317 (f2)
 hL2A1	$C_{66}H_{100}N_{16}O_{18}S$	3.5-5 (f1); 11.5-12.5 (f2); 12.5-13.5 (f3)	[M+H]= 1437.720 [M+Na]= 1459.702	1437.655; 1459.863 (f1) 1459.262 (f2) 1459.433 (f3)
 hL2A2	$C_{66}H_{100}N_{16}O_{18}S$	6.5-7.5	[M+H]= 1437.720 [M+Na]= 1459.702	1458.747
 hL2A3	$C_{66}H_{100}N_{16}O_{18}S$	30.5-31 (f1); 31-33.5 (f2)	[M+H]= 1437.720 [M+Na]= 1459.702	1436.878; 1458.864 (f1) 1437.200; 1458.190 (f2)
 hL2A4	$C_{71}H_{103}N_{15}O_{18}S$	42-42.5 (f1); 42.5-44 (f2)	[M+H]= 1486.740 [M+Na]= 1508.722	1487.214; 1508.178 (f1) 1486.406; 1508.364 (f2)
 hL2A5	$C_{70}H_{100}N_{16}O_{19}S$	39-41.5	[M+H]= 1501.715 [M+Na]= 1523.697	1501.307; 1523.288

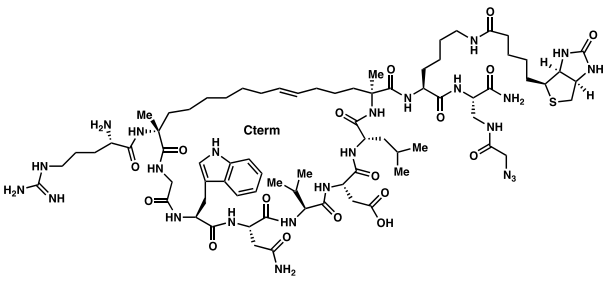
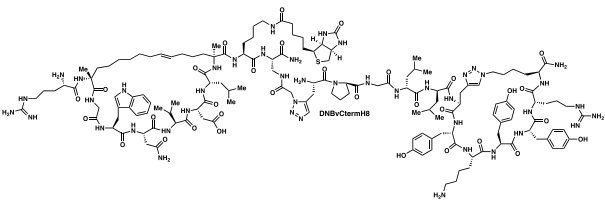
Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
	$C_{71}H_{103}N_{15}O_{18}S$	*	[M+H]= 1485.733	*
	$C_{68}H_{97}N_{15}O_{18}S$	35-38.5 (f1); 38.5-40.5 (f2); 40.5-42.5 (f3)	[M+H]= 1444.693 [M+Na]= 1466.675 [M+K]= 1482.649	1444.662; 1466.648 1482.232 (f1) 1443.219; 1465.181 (f2) 1443.628; 1465.609 (f3)
	$C_{72}H_{105}N_{15}O_{18}S$	39-42 (f1); 42-43 (f2); 43-44.5 (f3)	[M+H]= 1500.756 [M+Na]= 1522.738 [M+K]= 1538.712	1499.292; 1537.254 (f1) 1499.644; 1521.604 (f2) 1499.894; 1521.845 (f3)
	$C_{65}H_{99}N_{15}O_{17}S$	36.5-40 (f1); 40-42.5 (f2)	[M+H]= 1394.714 [M+Na]= 1416.696 [M+K]= 1432.670	1394.717; 1432.688 (f1) 1394.798; 1416.772 (f2)
	$C_{65}H_{99}N_{15}O_{17}S$	40-42.5 (f1); 42.5-45 (f2)	[M+H]= 1394.714 [M+Na]= 1416.696	1393.923; 1415.884 (f1) 1394.117 (f2)
	$C_{65}H_{99}N_{15}O_{17}S$	39.5-41 (f1); 41-42 (f2)	[M+H]= 1394.714 [M+Na]= 1416.696	1393.788; 1415.760 (f1) 1394.019; 1416.017 (f2)
	$C_{70}H_{107}N_{15}O_{19}S$	45.5-48.5 (f1); 51-55.3 (f2)	[M+H]= 1494.767 [M+Na]= 1516.749 [M+K]= 1532.723 [M-H <sub>2</sub> O+H]= 1476.756	1494.693; 1516.670; 1532.633 (f1) 1476.902 (f2)



Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
	$C_{71}H_{109}N_{15}O_{19}S$	45-47.5	[M+H]= 1509.787	1507.978
	$C_{72}H_{103}FN_{16}O_{18}S$	42-43 (f1); 43-45 (f2)	[M+H]= 1531.742 [M+Na]= 1553.724	1530.483; 1552.463 (f1) 1530.693; 1552.664 (f2)
	$C_{72}H_{102}F_2N_{16}O_{18}S$	40-42 (f1); 42-43.5 (f2); 43.5-44.5 (f3)	[M+H]= 1549.733 [M+Na]= 1571.714	1564.202 (f1) 1548.644; 1570.611 (f2) 1548.868; 1570.840 (f3)
	$C_{72}H_{103}FN_{16}O_{18}S$	43.5-44.5 (f1); 44.5-45.5 (f2)	[M+H]= 1531.742 [M+Na]= 1553.724	1530.534; 1552.511 (f1) 1530.800; 1552.775 (f2)
	$C_{72}H_{102}F_2N_{16}O_{18}S$	41.5-44.5 (f1); 44.5-45.5 (f2); 45.5-46.5 (f3)	[M+H]= 1549.733 [M+Na]= 1571.714	1564.199 (f1) 1548.644; 1570.611 (f2) 1576.669 (f3)
	$C_{74}H_{109}N_{15}O_{18}S$	43-45	[M+H]= 1528.787 [M+Na]= 1550.718	1527.260; 1550.217

Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
	$C_{71}H_{102}FN_{15}O_{17}S$	42-45 (f1); 45-46.5 (f2)	[M+H]= 1488.736 [M+Na]= 1510.718 [M+K]= 1526.692	1524.900 (f1) 1487.261; 1509.233 (f2)
	$C_{71}H_{101}F_2N_{15}O_{17}S$	46-47.5	[M+H]= 1506.727 [M+Na]= 1528.709	1506.250; 1527.212
	$C_{71}H_{102}FN_{15}O_{17}S$	44-45.5 (f1); 45.5-47 (f2)	[M+H]= 1488.736 [M+Na]= 1510.718 [M+K]= 1526.692	1487.286; 1509.237 (f1) 1487.513; 1509.467 (f2)
	$C_{71}H_{101}F_2N_{15}O_{17}S$	46-47.5 (f1); 47.5-49.5 (f2)	[M+H]= 1506.727 [M+Na]= 1528.709	1505.647; 1527.611 (f1) 1505.927 (f2)
	$C_{71}H_{102}FN_{15}O_{17}S$	46-47.5	[M+H]= 1488.736 [M+Na]= 1510.718	1489.199; 1509.170
	$C_{71}H_{101}F_2N_{15}O_{17}S$	42-46 (f1); 46-49.5 (f2)	[M+H]= 1506.727 [M+K]= 1544.683	1544.018 (f1) 1544.363 (f2)
	$C_{82}H_{125}N_{13}O_{19}S$	49-50 (f1); 50-50.5 (f2); 50.5-51 (f3)	[M+H]= 1628.901 [M+Na]= 1650.883	1649.923 (f1) 1648.839 (f2) 1649.211 (f3)
	$C_{82}H_{125}N_{13}O_{19}S$	48.5-49.5 (f1); 50-50.5 (f2); 50.5-51 (f3)	[M+H]= 1628.901 [M+Na]= 1650.883	1648.836 (f1) 1649.201 (f2) 1649.211 (f3)

Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
 <p>Ak-hL3V2G6</p>	$C_{82}H_{125}N_{13}O_{19}S$	48.5-49.5 (f1); 50-51 (f2)	[M+H]= 1628.901 [M+Na]= 1650.883	1651.204 (f1) 1650.208 (f2)
 <p>Ak-hL3V2G6</p>	$C_{82}H_{125}N_{13}O_{19}S$	34.5-35 (f1); 35-36.5 (f2)	[M+H]= 1628.901 [M+Na]= 1650.883	1650.626 (f1) 1649.784 (f2)
 <p>Tz-hL3V2G6</p>	$C_{76}H_{112}N_{16}O_{19}S$	42-44	[M+H]= 1585.809 [M+Na]= 1607.791	1584.127; 1606.108
 <p>Tz-hL3G1V3</p>	$C_{76}H_{112}N_{16}O_{19}S$	17-18 (f1); 18-18.5 (f2)	[M+H]= 1585.809 [M+Na]= 1607.791	1584.743; 1606.728 (f1) 1606.855 (f2)
 <p>Tz-hL3V2G8 heterodimer</p>	$C_{152}H_{224}N_{32}O_{38}S_2$	44-46	[M+Na]= 3192.592	3190.337
 <p>Tz-hL3G1V3 heterodimer</p>	$C_{152}H_{224}N_{32}O_{38}S_2$	19-20	[M+Na]= 3192.252	3191.441

Structure	Chemical Formula	HPLC Ret. Time (min)	Exp. mass (amu)	Obs. mass (amu)
	$C_{76}H_{120}N_{24}O_{17}S$	36.5-38 (f1); 38-39 (f2)	[M+H]= 1673.906 [M-N <sub>2</sub> +H]= 1647.916	1645.33; 1671.71 (f1) 1672.42 (f2)
	$C_{150}H_{226}N_{44}O_{32}S$	44-45.5 (f1); 45.5-46.5 (f2)	[M+H]= 3188.721	3187.95 (f1) 3190.21 (f2)

\*Spectra lost due to computer failure

## References

- (1) Expected masses were calculated using the mass calculator at the following:  
<http://www.lfd.uci.edu/~gohlke/molmass/?q=C152H224N32O38S2Na>.